





---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

NOVEL MOLECULES OF THE CARD-RELATED  
PROTEIN FAMILY AND USES THEREOF

Background of the Invention

5           In multicellular organisms, homeostasis is maintained by balancing the rate of cell proliferation against the rate of cell death. Cell proliferation is influenced by numerous growth factors and the expression of proto-oncogenes, which typically encourage progression through the cell cycle. In contrast, numerous events, including the expression of tumor suppressor genes, can lead to  
10 an arrest of cellular proliferation.

          In differentiated cells, a particular type of cell death called apoptosis occurs when an internal suicide program is activated. This program can be initiated by a variety of external signals as well as signals that are generated within the cell in response to, for example, genetic damage. For many years, the  
15 magnitude of apoptotic cell death was not appreciated because the dying cells are quickly eliminated by phagocytes, without an inflammatory response.

          The mechanisms that mediate apoptosis have been intensively studied. These mechanisms involve the activation of endogenous proteases, loss of mitochondrial function, and structural changes such as disruption of the  
20 cytoskeleton, cell shrinkage, membrane blebbing, and nuclear condensation due to degradation of DNA. The various signals that trigger apoptosis are thought to bring about these events by converging on a common cell death pathway that is regulated by the expression of genes that are highly conserved from worms, such as *C. elegans*, to humans. In fact, invertebrate model systems have been  
25 invaluable tools in identifying and characterizing the genes that control apoptosis. Through the study of invertebrates and more evolved animals, numerous genes that are associated with cell death have been identified, but the way in which their products interact to execute the apoptotic program is poorly understood.

          Caspases, a class of proteins central to the apoptotic program, are  
30 responsible for the degradation of cellular proteins that leads to the morphological changes seen in cells undergoing apoptosis. Caspases are cysteine

proteases having specificity for aspartate at the substrate cleavage site. An effector caspase is activated by an initiator caspase which cleaves the effector caspase at specific internal aspartate residues resulting in the separation of the large and small subunits of the effector caspase. For example, one of the

5 caspases identified in humans was previously known as the interleukin-1 $\alpha$  (IL-1 $\alpha$ ) converting enzyme (ICE), a cysteine protease responsible for the processing of pro-IL-1 $\alpha$  to the active cytokine. Overexpression of ICE in Rat-1 fibroblasts induces apoptosis (Miura et al., Cell 75:653, 1993).

Many caspases and proteins that interact with caspases possess

10 domains of about 60 amino acids called a caspase recruitment domain (CARD). Hofmann et al. (TIBS 22:155, 1997) and others have postulated that certain apoptotic proteins bind to each other via their CARDS and that different subtypes of CARDS may confer binding specificity, regulating the activity of various caspases, for example. The functional significance of CARDS have been

15 repeatedly demonstrated. For example, Duan et al. (Nature 385:86, 1997) showed that deleting the CARD at the N-terminus of RAIDD abolished the ability of RAIDD to bind to caspases.

Caspase-9 activation may precede the activation of all other cell death-related caspases in the mitochondrial pathways of apoptosis (See et al., J. Cell Biol. 144:281-292, 1999). Inactive procaspase-9 is activated by interaction

20 with a complex which includes Apaf-1, a CARD-containing protein, and other factors (Li et al., Cell 91:479, 1997; Srinivasula et al., Mol. Cell 1:949-959, 1998). Recognition of procaspase-9 by Apaf-1 occurs primarily through the interaction of the CARD of Apaf-1 with the prodomain of caspase-9. The CARD

25 of Apaf-1 shares about 20% sequence identity with the prodomain of procaspase-9. The prodomain of caspase-9 is a member of the CARD family of apoptotic signalling motifs (Hofmann and Bucher, Trends in Biochem. Sci. 22:155-156, 1997). A similar domain is present in caspase activating proteins CED-4 and RAIDD/CRADD as well as in initiator caspases CED-3 and caspase-2/ICH-1

30 (Duan and Dixit, Nature 385:86-89, 1997; Ahmad et al., Cancer Res. 57:615-619.



1997; Alnemri et al., Cell 87:171, 1996). Apaf-1 can bind several other caspases, e.g., caspase-4 and caspase-8 (Inohara et al., J. Biol. Chem. 273:12296-12300, 1998).

Nuclear factor- $\kappa$ B (NF- $\kappa$ B) is a transcription factor expressed in many cell types and which activates homologous or heterologous genes that have  $\kappa$ B sites in their promoters. Molecules that regulate NF- $\kappa$ B activation play a critical role in both apoptosis and inflammation. Quiescent NF- $\kappa$ B resides in the cytoplasm as a heterodimer of proteins referred to as p50 and p65 and is complexed with the regulatory protein I $\kappa$ B. NF- $\kappa$ B binding to I $\kappa$ B causes NF- $\kappa$ B to remain in the cytoplasm. At least two dozen stimuli that activate NF- $\kappa$ B are known (New England Journal of Medicine 336:1066, 1997) and they include cytokines, protein kinase C activators, oxidants, viruses, and immune system stimuli. NF- $\kappa$ B activating stimuli activate specific I $\kappa$ B kinases that phosphorylate I $\kappa$ B leading to its degradation. Once liberated from I $\kappa$ B, NF- $\kappa$ B translocates to the nucleus and activates genes with  $\kappa$ B sites in their promoters. The proinflammatory cytokines TNF- $\alpha$  and IL-1 induce NF- $\kappa$ B activation by binding their cell-surface receptors and activating the NF- $\kappa$ B-inducing kinase, NIK, and NF- $\kappa$ B. NIK phosphorylates the I $\kappa$ B kinases  $\alpha$  and  $\beta$  which phosphorylate I $\kappa$ B, leading to its degradation.

NF- $\kappa$ B and the NF- $\kappa$ B pathway has been implicated in mediating chronic inflammation in inflammatory diseases such as asthma, ulcerative colitis, rheumatoid arthritis (Epstein, New England Journal of Medicine 336:1066, 1997) and inhibiting NF- $\kappa$ B or NF- $\kappa$ B pathways may be an effective way of treating these diseases. NF- $\kappa$ B and the NF- $\kappa$ B pathway has also been implicated in atherosclerosis (Navab et al., American Journal of Cardiology 76:18C, 1995), especially in mediating fatty streak formation, and inhibiting NF- $\kappa$ B or NF- $\kappa$ B pathways may be an effective therapy for atherosclerosis. Among the genes activated by NF- $\kappa$ B are cIAP-1, cIAP-2, TRAF1, and TRAF2, all of which have been shown to protect cells from TNF- $\alpha$  induced cell death (Wang et al., Science 281:1680-83, 1998). CLAP, a protein which includes a CARD, activates the

Apaf-1-caspase-9 pathway and activates NF- $\kappa$ B by acting upstream of NIK and I $\kappa$ B kinase (Srinivasula et al., supra).

Bcl-2 family proteins are important regulators of pathways involved in apoptosis and can act to inhibit or promote cell death. Expression of certain anti-apoptotic Bcl-2 family members is commonly altered in cancerous cells, suppressing programmed cell death and extending tumor growth. Among the anti-apoptotic Bcl-2 family members thus far identified are Boo, Bcl-2, Bcl-x<sub>L</sub>, Bcl-w, NR-13, A1, and Mcl-2. Pro-apoptotic Bcl-2 family members include Bax, Bak, Bad, Bik, Bid, Hrk, Bim, and Bok/Mtd. Significantly, the anti-apoptotic Bcl-2 family member, Bcl-x<sub>L</sub>, has been shown to interact with Apaf-1 and block Apaf-1-dependent caspase-9 activation (Hu et al., Proc. Nat'l. Acad. Sci. 95:4386-4391, 1998). Boo, another anti-apoptotic Bcl-2 family member, interacts with Apaf-1 and caspase-9. Bak and Bik, pro-apoptotic Bcl-2 family members, can disrupt the association of Boo with Apaf-1 (Song et al., EMBO J. 18:167-178, 1999). Boo is thought to be involved in the control of ovarian atresia and sperm maturation. Diva, another member of the Bcl-2 family, inhibits binding of Bcl-x<sub>L</sub> to Apf-1, preventing Bcl-x<sub>L</sub> from binding to Apaf-1.

Neurotrophins (e.g., NGF), which are best known as neuronal survival factors, can mediate apoptosis via the p75 neurotrophin receptor (p75<sup>NTR</sup>). It is thought that p75<sup>NTR</sup> activation can lead to NF- $\kappa$ B activation (Carter et al., Science 272:542-545, 1996). It has been proposed that p75<sup>NTR</sup>-mediated cell death acts to ensure rapid cell death when a neuron is unable to obtain sufficient neurotrophins. This mechanism could, for example, cause the elimination of neurons that reach an inappropriate target or that reach an appropriate target at an inappropriate time (Miller and Kaplan, Cell Death and Diff. 5:343-345, 1998).

### Summary of the Invention

The present invention is based, at least in part, on the discovery of genes encoding CARD-3, CARD-4, CARD-5, and CARD-6. A full-length human CARD-3 cDNA is presented. Several CARD-4 cDNAs are presented.

- 5 Briefly, the CARD-4 gene can express a long transcript that encodes CARD-4L, a short transcript that encodes partial CARD-4S, or two CARD-4 splice variants (CARD-4Y and CARD-4Z). A full length cDNA sequence for the murine ortholog of CARD-4L is also presented. Full-length cDNAs encoding murine and human CARD-5 are presented. In addition, full-length cDNAs encoding
- 10 human and rat CARD-6 are presented.

CARD-3, CARD-4, CARD-5, and CARD-6 are intracellular proteins that are predicted to be involved in regulating caspase activation. CARD-4 is found to activate the NF- $\kappa$ B pathway and to enhance caspase 9-mediated cell death. In addition, proteins that bind to CARD-4 are presented including CARD-

15 3 and hNUDC.

The CARD-3 cDNA described below (SEQ ID NO:1) has a 1620 open reading frame (nucleotides 214 to 1833 of SEQ ID NO:1; SEQ ID NO:3) which encodes a 540 amino acid protein (SEQ ID NO:2). CARD-3 contains a kinase domain which extends from amino acid 1 to amino acid 300 of SEQ ID

20 NO:2; SEQ ID NO:4, followed by a linker domain at amino acid 301 to amino acid 431 of SEQ ID NO:2; SEQ ID NO:5 and a CARD at amino acid 432 to amino acid 540 of SEQ ID NO:2; SEQ ID NO:6.

At least four forms of CARD-4 exist in the cell, a long form, CARD-4L, a short form, CARD-4S, and two splice variants, CARD-4Y and CARD-4Z.

25 The cDNA of CARD-4L described below (SEQ ID NO:7) has a 2859 nucleotide open reading frame (nucleotides 245-3103 of SEQ ID NO:7; SEQ ID NO:9) which encodes a 953 amino acid protein (SEQ ID NO:8). CARD-4L protein possesses a CARD domain (amino acids 15-114; SEQ ID NO:10). The nucleotide sequence of the full length cDNA corresponding to the murine

30 ortholog of human CARD-4L is presented (SEQ ID NO:42) as is the predicted

amino acid sequence of murine CARD-4L (SEQ ID NO:43). A comparison between the predicted amino acid sequences of human CARD-4L and murine CARD-4L is also depicted in Figure 17.

Human CARD-4L is also predicted to have a nucleotide binding  
5 domain which extends from about amino acid 198 to about amino acid 397 of  
SEQ ID NO:8; SEQ ID NO:11, a Walker Box "A", which extends from about  
amino acid 202 to about amino acid 209 of SEQ ID NO:8; SEQ ID NO:12, a  
Walker Box "B", which extends from about amino acid 280 to about amino acid  
284, of SEQ ID NO:8; SEQ ID NO:13, a kinase 1a (P-loop) subdomain, which  
10 extends from about amino acid 127 to about amino acid 212 of SEQ ID NO:8;  
SEQ ID NO:46, a kinase 2 subdomain, which extends from about amino acid 273  
to about amino acid 288 of SEQ ID NO:8; SEQ ID NO:47, a kinase 3a  
subdomain, which extends from about amino acid 327 to about amino acid 338 of  
SEQ ID NO:8; SEQ ID NO:14, and ten Leucine-rich repeats which extend from  
15 about amino acid 674 to about amino acid 950 of SEQ ID NO:8. The first  
Leucine-rich repeat extends from about amino acid 674 to about amino acid 701  
of SEQ ID NO:8; SEQ ID NO:15. The second Leucine-rich repeat extends from  
about amino acid 702 to about amino acid 727 of SEQ ID NO:8; SEQ ID NO:16.  
The third Leucine-rich repeat extends from about amino acid 728 to about amino  
20 acid 754 of SEQ ID NO:8; SEQ ID NO:17. The fourth Leucine-rich repeat  
extends from about amino acid 755 to about amino acid 782 of SEQ ID NO:8;  
SEQ ID NO:18. The fifth Leucine-rich repeat extends from about amino acid  
783 to about amino acid 810 of SEQ ID NO:8; SEQ ID NO:19. The sixth  
Leucine-rich repeat extends from about amino acid 811 to about amino acid 838  
25 of SEQ ID NO:8; SEQ ID NO:20. The seventh Leucine-rich repeat extends from  
about amino acid 839 to about amino acid 866 of SEQ ID NO:8; SEQ ID NO:21.  
The eighth Leucine-rich repeat extends from about amino acid 867 to about  
amino acid 894 of SEQ ID NO:8; SEQ ID NO:22. The ninth Leucine-rich repeat  
extends from about amino acid 895 to about amino acid 922 of SEQ ID NO:8;  
30 SEQ ID NO:23 and the tenth leucine-rich repeat extends from about amino acid

923 to about amino acid 950 of SEQ ID NO:8; SEQ ID NO:24.

The partial cDNA of CARD-4S described below (SEQ ID NO:25) has a 1470 nucleotide open reading frame (nucleotides 1-1470 of SEQ ID NO:25; SEQ ID NO:27) which encodes a 490 amino acid protein (SEQ ID NO:26).

- 5 CARD-4S protein possesses a CARD domain (amino acids 1-74 of SEQ ID NO:26; SEQ ID NO:28). CARD-4S is predicted to have a P-Loop which extends from about amino acid 163 to about amino acid 170 of SEQ ID NO:26; SEQ ID NO:29, and a Walker Box "B" which extends from about amino acid 241 to about amino acid 245 of SEQ ID NO:26; SEQ ID NO:30.

- 10 A human CARD-4Y nucleotide cDNA sequence is presented (SEQ ID NO:38) as is the amino acid sequence of the predicted CARD-4Y product (SEQ ID NO:39). A human CARD-4Z nucleotide cDNA sequence is presented (SEQ ID NO:40) as is the amino acid sequence of the predicted CARD-4Z product (SEQ ID NO:41). A comparison of the CARD-4Y, CARD-4Z, and human  
15 CARD-4L predicted amino acid sequences is also shown in Figure 14.

- The 761 nucleotide murine CARD-5 cDNA described below (SEQ ID NO:60) has a 579 nucleotide open reading frame (nucleotides 89 to 668 of SEQ ID NO:60; SEQ ID NO:62) which encodes a 193 amino acid protein (SEQ ID NO:61). Murine CARD-5 contains a CARD domain which extends from amino  
20 acid 110 to amino acid 179 of SEQ ID NO:61 (SEQ ID NO:66).

- The 740 nucleotide human CARD-5 cDNA described below (SEQ ID NO:48) has a 585 nucleotide open reading frame (nucleotides 54 to 639 of SEQ ID NO:48; SEQ ID NO:50) which encodes a 195 amino acid protein (SEQ ID NO:49). Human CARD-5 contains a CARD domain which extends from amino  
25 acid 111 to amino acid 181 of SEQ ID NO:49 (SEQ ID NO:58).

- The 5252 nucleotide rat CARD-6 cDNA described below (SEQ ID NO:51) has a 2715 nucleotide open reading frame (nucleotides 169 to 2883 of SEQ ID NO:51; SEQ ID NO:53) which encodes a 905 amino acid protein (SEQ ID NO:52). Rat CARD-6 contains a CARD domain which extends from amino  
30 acid 1 to amino acid 108 of SEQ ID NO:52 (SEQ ID NO:59). Rat CARD-6 also

has a proline-rich C-terminus which extends from amino acid 698 to amino acid 905 of SEQ ID NO:52 (SEQ ID NO:65). This proline-rich domain includes five putative SH3 binding sites. These binding sites have the sequence PXXP and are located at amino acids 710 to 713 (PAHP), 806 to 809 (PLRP), 819 to 822 (PIPP), 857 to 860 (PPHP), and 881 to 884 (PSQP) of SEQ ID NO:52.

The 4244 human CARD-6 cDNA described below (SEQ ID NO:54) has a 3111 nucleotide open reading frame (nucleotides 200 to 3310 of SEQ ID NO:54; SEQ ID NO:56) which encodes a 1037 amino acid protein (SEQ ID NO:55). Human CARD-6 includes a CARD domain which extends from amino acid 5 to amino acid 92 of SEQ ID NO:55 (SEQ ID NO:64).

Like other proteins containing a CARD domain, CARD-3, CARD-4, CARD-5, and CARD-6 to participate in the network of interactions that lead to caspase activity. Human CARD-4L likely plays a functional role in caspase activation similar to that of Apaf-1 (Zou et al. (1997) Cell 90:405-413). For example, upon activation, CARD-4L binds a nucleotide, thus allowing CARD-4L to bind and activate a CARD-containing caspase via a CARD-CARD interaction, leading to apoptotic death of the cell. CARD-3, CARD-4, CARD-5, and CARD-6 molecules are useful as modulating agents in regulating a variety of cellular processes including cell growth and cell death. In one aspect, this invention provides isolated nucleic acid molecules encoding CARD-3, CARD-4, CARD-5, or CARD-6 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of CARD-3, CARD-4, CARD-5, or CARD-6 encoding nucleic acids.

The invention encompasses methods of diagnosing and treating patients who are suffering from a disorder associated with an abnormal level or rate (undesirably high or undesirably low) of apoptotic cell death, abnormal activity of the Fas/APO-1 receptor complex, abnormal activity of the TNF receptor complex, or abnormal activity of a caspase by administering a compound that modulates the expression of CARD-3, CARD-4, CARD-5, or CARD-6 (at the DNA, mRNA or protein level, e.g., by altering mRNA splicing)

or by altering the activity of CARD-3, CARD-4, CARD-5, or CARD-6. Examples of such compounds include small molecules, antisense nucleic acid molecules, ribozymes, and polypeptides.

Certain disorders are associated with an increased number of  
5 surviving cells, which are produced and continue to survive or proliferate when apoptosis is inhibited or occurs at an undesirably low rate. Compounds that modulate the expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 can be used to treat or diagnose such disorders. These disorders include cancer (particularly follicular lymphomas, chronic myelogenous leukemia, melanoma,  
10 colon cancer, lung carcinoma, carcinomas associated with mutations in p53, and hormone-dependent tumors such as breast cancer, prostate cancer, and ovarian cancer). Such compounds can also be used to treat viral infections (such as those caused by herpesviruses, poxviruses, and adenoviruses). Failure to remove autoimmune cells that arise during development or that develop as a result of  
15 somatic mutation during an immune response can result in autoimmune disease. Thus, autoimmune disorders can be caused by an undesirably low levels of apoptosis. Accordingly, modulators of CARD-3, CARD-4, CARD-5, or CARD-6 activity or expression can be used to treat autoimmune disorders (e.g., systemic lupus erythematosus, immune-mediated glomerulonephritis, and arthritis).

20 Many diseases are associated with an undesirably high rate of apoptosis. Modulators of CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity can be used to treat or diagnose such disorders. For example, populations of cells are often depleted in the event of viral infection, with perhaps the most dramatic example being the cell depletion caused by the human  
25 immunodeficiency virus (HIV). Surprisingly, most T cells that die during HIV infections do not appear to be infected with HIV. Although a number of explanations have been proposed, recent evidence suggests that stimulation of the CD4 receptor results in the enhanced susceptibility of uninfected T cells to undergo apoptosis. A wide variety of neurological diseases are characterized by  
30 the gradual loss of specific sets of neurons. Such disorders include Alzheimer's

disease, Parkinson's disease, amyotrophic lateral sclerosis (ALS) retinitis pigmentosa, spinal muscular atrophy, and various forms of cerebellar degeneration. The cell loss in these diseases does not induce an inflammatory response, and apoptosis appears to be the mechanism of cell death. In addition, a number of hematologic diseases are associated with a decreased production of blood cells. These disorders include anemia associated with chronic disease, aplastic anemia, chronic neutropenia, and the myelodysplastic syndromes. Disorders of blood cell production, such as myelodysplastic syndrome and some forms of aplastic anemia, are associated with increased apoptotic cell death within the bone marrow. These disorders could result from the activation of genes that promote apoptosis, acquired deficiencies in stromal cells or hematopoietic survival factors, or the direct effects of toxins and mediators of immune responses. Two common disorders associated with cell death are myocardial infarctions and stroke. In both disorders, cells within the central area of ischemia, which is produced in the event of acute loss of blood flow, appear to die rapidly as a result of necrosis. However, outside the central ischemic zone, cells die over a more protracted time period and morphologically appear to die by apoptosis.

Proteins containing a CARD domain are thought to be involved in various inflammatory disorders. Accordingly, CARD-3, CARD-4, CARD-5, and CARD-6 polypeptides, nucleic acids and modulators of CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity can be used to treat immune disorders. Such immune disorders include, but are not limited to, chronic inflammatory diseases and disorders, such as Crohn's disease, reactive arthritis, including Lyme disease, insulin-dependent diabetes, organ-specific autoimmunity, including multiple sclerosis, Hashimoto's thyroiditis and Grave's disease, contact dermatitis, psoriasis, graft rejection, graft versus host disease, sarcoidosis, atopic conditions, such as asthma and allergy, including allergic rhinitis, gastrointestinal allergies, including food allergies, eosinophilia, conjunctivitis, glomerular nephritis, certain pathogen susceptibilities such as



helminthic (e.g., leishmaniasis), certain viral infections, including HIV, and bacterial infections, including tuberculosis and lepromatous leprosy.

In addition to the aforementioned disorders, CARD-3, CARD-4, CARD-5, and CARD-6 polypeptides, nucleic acids, and modulators of CARD-3, CARD-4, CARD-5 or CARD-6 expression or activity can be used to treat disorders of cell signaling and disorders of tissues in which CARD-3, CARD-4, CARD-5 or CARD-6 is expressed.

The invention features a nucleic acid molecule which is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62. the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number 203037 (the "cDNA of ATCC 203037"), the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number 203035 (the "cDNA of ATCC 203035"), the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number 203036 (the "cDNA of ATCC 203036"), the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number PTA-211 (the "cDNA of ATCC PTA-211"), the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number PTA-212 ("the cDNA of ATCC PTA-212"), the nucleotide sequence of the cDNA insert of the plasmid deposited with the ATCC as Accession Number PTA-213 (the "cDNA of ATCC PTA-213"), or a complement thereof.

The invention features a nucleic acid molecule which includes a fragment of at least 150 (300, 325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1300, 1600 or 1931) nucleotides of the nucleotide sequence shown in SEQ ID NO:1, or SEQ ID NO:3, or the nucleotide sequence of the cDNA ATCC 203037, or a complement thereof.

The invention also features a nucleic acid molecule which includes a fragment of at least 150 (350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1300, 1600, 1900, 2100, 2400, 2700, 3000, or 3382) nucleotides of the nucleotide sequence shown in SEQ ID NO:7, SEQ ID NO:9, or the nucleotide sequence of  
5 the cDNA ATCC 203035, or a complement thereof.

Also within the invention is a nucleic acid molecule which includes a fragment of at least 150 (350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1300, 1600, 1900, 2100, 2400, 2700, and 3080) nucleotides of the nucleotide sequence shown in SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID  
10 NO:40, or the nucleotide sequence of the cDNA of ATCC 203036, or a complement thereof.

The invention also features a nucleic acid molecule which includes a fragment of at least 150 (350, 400, 450, 500, 550, 600, 650, 700, and 761) nucleotides of the nucleotide sequence shown in SEQ ID NO:60, SEQ ID NO:62,  
15 or the nucleotide sequence of the cDNA of ATCC PTA-212, or a complement thereof.

The invention also features a nucleic acid molecule which includes a fragment of at least 150 (350, 400, 450, 500, 550, 600, 650, 700, and 740) nucleotides of the nucleotide sequence shown in SEQ ID NO:48, SEQ ID NO:50,  
20 the cDNA of ATCC PTA-213, or a complement thereof.

The invention also features a nucleic acid molecule which includes a fragment of at least 150 (350, 400, 450, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, and 5252) nucleotides of the nucleotide sequence shown in SEQ ID NO:51, SEQ ID NO:53, or a complement  
25 thereof.

The invention also features a nucleic acid molecule which includes a fragment of at least 150 (200, 300, 400, 500, 600, 700, 800, 900, 1000, 1400, 1800, 2200, 2600, or 3000) nucleotides of the nucleotide sequence shown in SEQ ID NO:54, SEQ ID NO:56, the cDNA of ATCC PTA-213, or a complement  
30 thereof.

The invention features a nucleic acid molecule which includes a nucleotide sequence encoding a protein having an amino acid sequence that is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:61, or the amino acid sequence encoded by the cDNA of ATCC 203037, the amino acid sequence encoded by the cDNA of ATCC 203035, the amino acid sequence encoded by the cDNA of ATCC 203036, the amino acid sequence encoded by the cDNA of ATCC PTA-211, the amino acid sequence encoded by the cDNA of ATCC PTA-212, or the amino acid sequence encoded by the cDNA of ATCC PTA-213.

In an embodiment, a CARD-3 nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:1, or SEQ ID NO:3, or the nucleotide sequence of the cDNA of ATCC 203037.

In another embodiment, a CARD-4L nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:7, or SEQ ID NO:9, or the nucleotide sequence of the cDNA of ATCC 203035.

In yet another embodiment, a CARD-4S nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:25, or SEQ ID NO:27, or the nucleotide sequence of the cDNA of ATCC 203036. In another embodiment, a murine CARD-4L nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:42.

In another embodiment, a CARD-4Y nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:38.

In another embodiment, a CARD-4Z nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:40.

In another embodiment, a human CARD-5 nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:48, SEQ ID NO:50 or the nucleotide sequence of the cDNA of ATCC PTA-213. In another embodiment, a murine CARD-5 nucleic acid molecule has the nucleotide sequence shown in

SEQ ID NO:60 or SEQ ID NO:62.

In yet another embodiment, a rat CARD-6 nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:51, SEQ ID NO:53, or the nucleotide sequence of the cDNA of ATCC PTA-211.

- 5 In still another embodiment, a human CARD-6 nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:54, SEQ ID NO:56, or the nucleotide sequence of the cDNA of ATCC PTA-213.

- Also within the invention is a nucleic acid molecule which encodes a fragment of a polypeptide having the amino acid sequence of SEQ ID NO:2, SEQ  
 10 ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:61, the fragment including at least 15 (25, 30, 50, 100, 150, 300, 400 or 540, 600, 700, 800, 900) contiguous amino acids of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, or  
 15 SEQ ID NO:55, SEQ ID NO:61, the polypeptide encoded by the cDNA of ATCC Accession Number 203037, the polypeptide encoded by the cDNA of ATCC Accession Number 203035, the polypeptide encoded by the cDNA of ATCC Accession Number 203036, the polypeptide encoded by the cDNA of ATCC Accession Number PTA-211, the polypeptide encoded by the cDNA of ATCC  
 20 Accession Number PTA-212, or the polypeptide encoded by the cDNA of ATCC Accession Number PTA-213.

- The invention includes a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ  
 25 ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, SEQ ID NO:61, or an amino acid sequence encoded by the cDNA of ATCC Accession Number 203037, 203035, 203036, PTA-211, PTA-212, or PTA-213, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ  
 30 ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID

NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of PTA-213 under stringent  
5 conditions.

In general, an allelic variant of a gene will be readily identifiable as mapping to the same chromosomal location as said gene. For example, in Example 6, the chromosomal location of the human CARD-4 gene is discovered to be chromosome 7 close to the SHGC-31928 genetic marker. Allelic variants  
10 of human CARD-4 will be readily identifiable as mapping to the human CARD-4 locus on chromosome 7 near genetic marker SHGC-31928.

Also within the invention are: an isolated CARD-3 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:2; an isolated CARD-3  
15 protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the kinase domain of SEQ ID NO:2 (e.g., about amino acid residues 1 to 300 of SEQ ID NO:2; SEQ ID NO:4); and an isolated CARD-3 protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the linker domain of SEQ ID NO:2 (e.g., about amino acid residues 301 to 431 of  
20 SEQ ID NO:2; SEQ ID NO:5); an isolated CARD-3 protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the CARD domain of SEQ ID NO:2 (e.g., about amino acid residues 432 to 540 of SEQ ID NO:2; SEQ ID NO:6); an isolated CARD-4L protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the  
25 amino acid sequence of SEQ ID NO:8; an isolated CARD-4L protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the CARD domain of SEQ ID NO:8 (e.g., about amino acid residues 15 to 114 of SEQ ID NO:8; SEQ ID NO:10); an isolated CARD-4L protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the nucleotide  
30 binding domain of SEQ ID NO:8 (e.g., about amino acid residues 198 to 397 of

SEQ ID NO:8; SEQ ID NO:11; an isolated CARD-4L protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the kinase 1a (P-loop) subdomain SEQ ID NO:8 (e.g., about amino acid 127 to about amino acid 212 of SEQ ID NO:8; SEQ ID NO:46); an isolated CARD-4L protein having

5 an amino acid sequence that is at least about 85%, 95%, or 98% identical to the kinase 2 subdomain of SEQ ID NO:8 (e.g., about amino acid 273 to about amino acid 288 of SEQ ID NO:8; SEQ ID NO:47); an isolated CARD-4L protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to a kinase 3a subdomain of SEQ ID NO:8 (e.g., about amino acid residues 327 to

10 338 of SEQ ID NO:8; SEQ ID NO:14); an isolated CARD-4L protein having an amino acid sequence that is at least about 85%, 95%, or 98% identical to the Leucine-rich repeats of SEQ ID NO:8 (e.g., about amino acid residues 674 to 701 of SEQ ID NO:8; SEQ ID NO:15; from amino acid 702 to amino acid 727 of SEQ ID NO:8; SEQ ID NO:16; which extends from amino acid 728 to amino

15 acid 754 SEQ ID NO:8; SEQ ID NO:17; from amino acid 755 to amino acid 782 of SEQ ID NO:8; SEQ ID NO:18; from amino acid 783 to amino acid 810 of SEQ ID NO:8; SEQ ID NO:19; from amino acid 811 to amino acid 838 of SEQ ID NO:8; SEQ ID NO:20 from amino acid 839 to amino acid 866 of SEQ ID NO:8; SEQ ID NO:21; from amino acid 867 to amino acid 894 of SEQ ID NO:8;

20 SEQ ID NO:22; from amino acid 895 to amino acid 922 of SEQ ID NO:8; SEQ ID NO:23; and from amino acid 923 to amino acid 950 of SEQ ID NO:8; SEQ ID NO:24); an isolated CARD-4S protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:26; an isolated CARD-4S protein having an amino acid

25 sequence that is at least about 85%, 95%, or 98% identical to the CARD domain of SEQ ID NO:26 (e.g., about amino acid residues 1 to 74 of SEQ ID NO:26; SEQ ID NO:28). Also within the invention are: an isolated murine CARD-4L protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:43. Also

30 within the invention are: an isolated CARD-4Y protein having an amino acid

sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:39. Also within the invention are: an isolated CARD-4Z protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of  
5 SEQ ID NO:41.

Also within the invention are: an isolated CARD-5 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:49 and an isolated CARD-5 protein comprising an amino acid sequence that is at least about 90%,  
10 95%, or 98% identical to SEQ ID NO:58 (CARD domain).

Also within the invention are an isolated CARD-5 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:60 and an isolated CARD-5 protein comprising an amino acid sequence that is at least about 90%,  
15 95%, or 98% identical to SEQ ID NO:57 (CARD domain).

The invention also includes: an isolated CARD-6 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:52 and an isolated CARD-6 protein having an amino acid sequence that is at least about 90%, 95%,  
20 or 98% identical to SEQ ID NO:59 (CARD domain).

The invention also includes: an isolated CARD-6 protein having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:55 and an isolated CARD-6 protein having an amino acid sequence that is at least about 90%, 95%,  
25 or 98% identical to SEQ ID NO:64 (CARD domain).

Also within the invention are: an isolated CARD-3 protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:3 or the cDNA of ATCC 203037; an isolated CARD-3 protein which is encoded by a nucleic  
30 acid molecule having a nucleotide sequence at least about 65% preferably 75%.

85%, or 95% identical to the kinase domain encoding portion of SEQ ID NO:1 (e.g., about nucleotides 213 to 1113 of SEQ ID NO:1); an isolated CARD-3 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 65% preferably 75%, 85%, or 95% identical the linker domain encoding portion of SEQ ID NO:1 (e.g., about nucleotides 1114 to 1506 of SEQ ID NO:1); and an isolated CARD-3 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 65% preferably 75%, 85%, or 95% identical the CARD domain encoding portion of SEQ ID NO:1 (e.g., about nucleotides 1507 to 1833 of SEQ ID NO:1); and an isolated CARD-3 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:3 or the non-coding strand of the cDNA of ATCC 203037. Also within the invention are: an isolated CARD-4Y protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:38. Also within the invention are nucleic acid molecules which include about nucleotides 2759 to 2842 of SEQ ID NO:7; about nucleotides 2843 to 2926 of SEQ ID NO:7; about nucleotides 2927 to 3010 of SEQ ID NO:7; about nucleotides 3011 to 3094 of SEQ ID NO:7; and an isolated CARD-4L protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:9, or the non-coding strand of the cDNA of ATCC 203035.

Also within the invention are an isolated CARD-4S protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:27; an isolated CARD-3 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 65% preferably 75%, 85%, or 95% identical the CARD domain encoding portion of SEQ ID NO:25 (e.g., about nucleotides 1 to 222 of SEQ ID NO:25); an isolated CARD-3 protein which is encoded by a



nucleic acid molecule having a nucleotide sequence at least about 65% preferably 75%, 85%, or 95% identical the P-Loop encoding portion of SEQ ID NO:25 (e.g., about nucleotides 485 to 510 of SEQ ID NO:25).

Also within the invention are an isolated CARD-5 protein which is  
5 encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:48 or the cDNA of ATCC PTA-213; an isolated CARD-5 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 90% preferably 95%, or 98% identical to the CARD encoding portion of SEQ ID NO:48 (e.g.,  
10 about nucleotides 383 to 596 of SEQ ID NO:48); and an isolated CARD-5 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:48 or the non-coding strand of the cDNA of ATCC PTA-213.

15 Also within the invention are an isolated CARD-5 protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:60; an isolated CARD-5 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 90% preferably 95%, or 98% identical to the  
20 CARD encoding portion of SEQ ID NO:60 (e.g., about nucleotides 416 to 625 of SEQ ID NO:60); and an isolated CARD-5 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:60.

25 Also within the invention are an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:51; an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 90% preferably 95%, or 98% identical to the  
30 CARD encoding portion of SEQ ID NO:51 (e.g., about nucleotides 169 to 456 of

SEQ ID NO:51); and an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:51.

5           Also within the invention are an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, preferably 75%, 85%, or 95% identical to SEQ ID NO:54; an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence at least about 90% preferably 95%, or 98% identical to the  
10   CARD encoding portion of SEQ ID NO:54; and an isolated CARD-6 protein which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:54.

          Another embodiment of the invention features CARD-3, CARD-4,  
15   CARD-5, or CARD-6 nucleic acid molecules which specifically detect CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules, relative to nucleic acid molecules encoding other members of the CARD superfamily. For example, in one embodiment, a CARD-4L nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ  
20   ID NO:7, SEQ ID NO:9, or the cDNA of ATCC 203035, or a complement thereof. In another embodiment, the CARD-4L nucleic acid molecule is at least 300 (350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1000, 1300, 1600, 1900, 2100, 2400, 2700, 3000, or 3382) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide  
25   sequence shown in SEQ ID NO:7, SEQ ID NO:9, the cDNA of ATCC 203035, or a complement thereof. In another embodiment, an isolated CARD-4L nucleic acid molecule comprises nucleotides 287 to 586 of SEQ ID NO:7, encoding the CARD domain of CARD-4L, or a complement thereof. In yet another embodiment, the invention provides an isolated nucleic acid molecule which is  
30   antisense to the coding strand of a CARD-4L nucleic acid.

In another embodiment, a CARD-5 nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:48, SEQ ID NO:50, or the cDNA of ATCC PTA-213, or a complement thereof. In another embodiment, the CARD-5 nucleic acid molecule is at least 300 (350, 400, 450, 500, 550, 585, 600, 650, 700, or 740) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:48, SEQ ID NO:50, the cDNA of ATCC PTA-213, or a complement thereof. In another embodiment, an isolated CARD-5 nucleic acid molecule comprises nucleotides 383 to 596 of SEQ ID NO:48, encoding the CARD of CARD-5. In yet another embodiment, the invention provides an isolated nucleic acid molecule which is antisense to the coding strand of a CARD-5 nucleic acid.

Another aspect of the invention provides a vector, e.g., a recombinant expression vector, comprising a CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecule of the invention. In another embodiment the invention provides a host cell containing such a vector. The invention also provides a method for producing CARD-3, CARD-4, CARD-5, or CARD-6 protein by culturing, in a suitable medium, a host cell of the invention containing a recombinant expression vector such that a CARD-3, CARD-4, CARD-5, or CARD-6 protein is produced.

Another aspect of this invention features isolated or recombinant CARD-3, CARD-4, CARD-5, or CARD-6 proteins and polypeptides. Preferred CARD-3, CARD-4, CARD-5, or CARD-6 proteins and polypeptides possess at least one biological activity possessed by naturally occurring human CARD-3, CARD-4, CARD-5, or CARD-6, e.g., (1) the ability to form protein:protein interactions with proteins in the apoptotic signaling pathway; (2) the ability to form CARD-CARD interactions with proteins in the apoptotic signaling pathway; (3) the ability to bind a CARD-3, CARD-4, CARD-5, or CARD-6 ligand; and (4) the ability to bind to an intracellular target. Other activities include: (1) modulation of cellular proliferation; (2) modulation of cellular differentiation; (3)

modulation of cellular death; and (4) modulation of the NF- $\kappa$ B pathway.

The CARD-3, CARD-4, CARD-5, or CARD-6 proteins of the present invention, or biologically active portions thereof, can be operatively linked to a non-CARD-3, non-CARD-4, non-CARD-5, or non-CARD-6 polypeptide (e.g.,  
5 heterologous amino acid sequences) to form CARD-3, CARD-4, CARD-5, or CARD-6 fusion proteins, respectively. The invention further features antibodies that specifically bind CARD-3, CARD-4, CARD-5, or CARD-6 proteins, such as monoclonal or polyclonal antibodies. In addition, the CARD-3, CARD-4, CARD-5, or CARD-6 proteins or biologically active portions thereof can be  
10 incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of CARD-3, CARD-4, CARD-5, or CARD-6 activity or expression in a biological sample by contacting the biological sample with an  
15 agent capable of detecting an indicator of CARD-3, CARD-4, CARD-5, or CARD-6 activity such that the presence of CARD-3, CARD-4, CARD-5, or CARD-6 activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating CARD-3, CARD-4, CARD-5, or CARD-6 activity comprising contacting a cell  
20 with an agent that modulates (inhibits or stimulates) CARD-3, CARD-4, CARD-5, or CARD-6 activity or expression such that CARD-3, CARD-4, CARD-5, or CARD-6 activity or expression in the cell is modulated. In one embodiment, the agent is an antibody that specifically binds to CARD-3, CARD-4, CARD-5, or CARD-6 protein. In another embodiment, the agent modulates expression of  
25 CARD-3, CARD-4, CARD-5, or CARD-6 by modulating transcription of a CARD-3, CARD-4, CARD-5, or CARD-6 gene, splicing of a CARD-3, CARD-4, CARD-5, or CARD-6 mRNA, or translation of a CARD-3, CARD-4, CARD-5, or CARD-6 mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of  
30 the CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or the CARD-3, CARD-4.

CARD-5, or CARD-6 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid expression or activity or related to  
5 CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity by administering an agent which is a CARD-3, CARD-4, CARD-5, or CARD-6 modulator to the subject. In one embodiment, the CARD-3, CARD-4, CARD-5, or CARD-6 modulator is a CARD-3, CARD-4, CARD-5, or CARD-6 protein. In another embodiment the CARD-3, CARD-4, CARD-5, or CARD-6 modulator is  
10 a CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecule. In other embodiments, the CARD-3, CARD-4, CARD-5, or CARD-6 modulator is a peptide, peptidomimetic, or other small molecule.

The present invention also provides a diagnostic assay for identifying the presence or absence of a genetic lesion or mutation characterized by at least  
15 one of: (i) aberrant modification or mutation of a gene encoding a CARD-3, CARD-4, CARD-5, or CARD-6 protein; (ii) mis-regulation of a gene encoding a CARD-3, CARD-4, CARD-5, or CARD-6 protein; (iii) aberrant RNA splicing; and (iv) aberrant post-translational modification of a CARD-3, CARD-4, CARD-5, or CARD-6 protein, wherein a wild-type form of the gene encodes a protein  
20 with a CARD-3, CARD-4, CARD-5, or CARD-6 activity.

In another aspect, the invention provides a method for identifying a compound that binds to or modulates the activity of a CARD-3, CARD-4, CARD-5, or CARD-6 protein. In general, such methods entail measuring a biological activity of a CARD-3, CARD-4, CARD-5, or CARD-6 protein in the  
25 presence and absence of a test compound and identifying those compounds which alter the activity of the CARD-3, CARD-4, CARD-5, or CARD-6 protein.

The invention also features methods for identifying a compound which modulates the expression of CARD-3, CARD-4, CARD-5, or CARD-6 by measuring the expression of CARD-3, CARD-4, CARD-5, or CARD-6 in the  
30 presence and absence of a compound.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

#### Brief Description of the Drawings

5           Figure 1 depicts the cDNA sequence (SEQ ID NO:1) of human CARD-3. The open reading frame of CARD-3 (SEQ ID NO:1) extends from nucleotide 213 to nucleotide 1833 nucleotide (SEQ ID NO:3).

          Figure 2 depicts the predicted amino acid sequence (SEQ ID NO:2) of human CARD-3.

10           Figure 3 depicts the cDNA sequence (SEQ ID NO:7) of CARD-4L. The open reading frame of SEQ ID NO:7 extends from nucleotide 245 to nucleotide 3103 (SEQ ID NO:9).

          Figure 4 depicts the predicted amino acid sequence (SEQ ID NO:8) of human CARD-4L.

15           Figure 5 depicts the partial cDNA sequence (SEQ ID NO:25) of CARD-4S and the predicted amino acid sequence (SEQ ID NO:25) of human CARD-4S. The open reading frame of CARD-4 (SEQ ID NO:25) extends from nucleotide 1 to nucleotide 1470 (SEQ ID NO:27).

          Figure 6 depicts the predicted amino acid sequence (SEQ ID NO:26) of human CARD-4S.

          Figure 7 depicts an alignment of the CARD domains of CARD-4 (SEQ ID NO:10), CARD-3 (SEQ ID NO:6), ARC-CARD (SEQ ID NO:31), cIAP1-CARD (SEQ ID NO:32), and cIAP2-CARD (SEQ ID NO:33).

25           Figure 8 is a plot showing predicted structural features of human CARD-4L.

          Figure 9 is a plot showing predicted structural features of human CARD-4S.

          Figure 10 depicts the cDNA sequence (SEQ ID NO:38) of the human CARD-4Y splice variant clone. The predicted open reading frame of the human  
30   CARD-4Y splice variant clone extends from nucleotide 438 to nucleotide 1184.

Figure 11 depicts the amino acid sequence (SEQ ID NO:39) of the protein predicted to be encoded by the human CARD-4Y cDNA open reading frame.

Figure 12 depicts the cDNA sequence (SEQ ID NO:40) of the human  
5 CARD-4Z splice variant clone. The predicted open reading frame of the human CARD-4Z splice variant clone extends from nucleotide 489 to nucleotide 980.

Figure 13 depicts the amino acid sequence (SEQ ID NO:41) of the protein predicted to be encoded by the human CARD-4Z cDNA open reading frame.

10 Figure 14 depicts an alignment of human CARD-4L (SEQ ID NO:8), the predicted amino acid sequence of human CARD-4Y (SEQ ID NO:39), and the predicted amino acid sequence of human CARD-4Z (SEQ ID NO:41).

Figure 15 depicts the nucleotide sequence of the murine CARD-4L cDNA (SEQ ID NO:42).

15 Figure 16 depicts the predicted amino acid sequence of murine CARD-4L (SEQ ID NO:43).

Figure 17 depicts an alignment of human CARD-4L (SEQ ID NO:8) and the predicted amino acid sequence of murine CARD-4L (SEQ ID NO:43).

Figure 18 depicts a 32042 nucleotide genomic sequence of CARD-4.

20 Figure 19 depicts the nucleotide sequence of a murine CARD-5 cDNA (SEQ ID NO:60). The open reading frame of this cDNA extends from nucleotide 89 to nucleotide 667 of SEQ ID NO:60 (SEQ ID NO:62) and encodes a 193 amino acid protein (SEQ ID NO:61).

Figure 20 depicts a hydropathy plot of murine CARD-5. Relatively  
25 hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 21 depicts the nucleotide sequence of a human CARD-5 cDNA  
30 (SEQ ID NO:48). The open reading frame of this cDNA extends from nucleotide

53 to nucleotide 638 of SEQ ID NO:48 (SEQ ID NO:50) and encodes a 195 amino acid protein SEQ ID NO:49).

Figure 22 depicts a hydropathy plot of human CARD-5. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 23 depicts an alignment of the cDNA sequences of murine CARD-5 (SEQ ID NO:60) and human CARD-5 (SEQ ID NO:48). This alignment was created using ALIGN (version 2.0; PAM120 scoring matrix; -12/-4 gap penalty). In this alignment the sequences are 68.2% identical.

Figure 24 depicts an alignment of the amino acid sequences of murine CARD-5 (SEQ ID NO:61) and human CARD-5 (SEQ ID NO:49). This alignment was created using ALIGN (version 2.0; PAM120 scoring matrix; -12/-4 gap penalty). In this alignment the sequences are 71.8% identical.

Figure 25 depicts the nucleotide sequence of a rat CARD-6 cDNA (SEQ ID NO:51). The open reading frame of this cDNA extends from nucleotide 169 to nucleotide 2883 of SEQ ID NO:51 (SEQ ID NO:53) and encodes a 505 amino acid protein (SEQ ID NO:52).

Figure 26 depicts a hydropathy plot of rat CARD-6. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 27 depicts an alignment of the CARD domains of murine CARD-5 (SEQ ID NO:57), human CARD-5 (SEQ ID NO:58), and RAIDD (SEQ ID NO:61).

Figure 28 depicts the nucleotide sequence of a human CARD-6 cDNA (SEQ ID NO:54). The open reading frame of this cDNA extends from nucleotide 200 to 3310 of SEQ ID NO:54 (SEQ ID NO:56) and encodes a 1037 amino acid



protein (SEQ ID NO:55).

Figure 29 depicts a hydropathy plot of human CARD-6. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line.

5           Figure 30 depicts an alignment of the CARD domain of human CARD-6 (SEQ ID NO:64) with a consensus CARD domain (SEQ ID NO:67). In this depiction of the consensus sequence, more conserved residues are indicated by uppercase letters and less conserved residues are indicated by lowercase letters.

10           Figure 31 depicts an alignment of the CARD domains of human CARD-3, human CARD-4, human CARD-5, murine CARD-5, human CARD-6, and rat CARD-6. This alignment was created using the Clustal method with PAM250 residue weight table. A consensus sequence is also depicted (SEQ ID NO:\_\_\_).

15

#### Detailed Description of the Invention

The present invention is based, in part, on the discovery of cDNA molecules encoding human CARD-3, human CARD-4, partial murine CARD-4I., murine CARD-5, human CARD-5, rat CARD-6, and human CARD-6 proteins.

20

TABLE 1

Summary of CARD-3, CARD-4, CARD-5, and CARD-6 Sequence Information.

5

Gene	cDNA	Protein	ORF	Figure	Accession Number
human CARD-3	SEQ ID NO:1	SEQ ID NO:2	SEQ ID NO:3	Figs. 1-2	203037
human CARD-4L	SEQ ID NO:7	SEQ ID NO:8	SEQ ID NO:9	Figs. 3-4	203035
human CARD-4S	SEQ ID NO:25	SEQ ID NO:26	SEQ ID NO:27	Figs. 5-6	203036
human CARD-4Y	SEQ ID NO:38	SEQ ID NO:39		Figs. 10-11	
human CARD-4Z	SEQ ID NO:40	SEQ ID NO:41		Figs. 12-13	
murine CARD-4L	SEQ ID NO:42	SEQ ID NO:43		Figs. 15-16	
human CARD-5	SEQ ID NO:48	SEQ ID NO:49	SEQ ID NO:50	Fig. 21	PTA-213
murine CARD-5	SEQ ID NO:60	SEQ ID NO:61	SEQ ID NO:62	Fig. 19	PTA-212
human CARD-6	SEQ ID NO:54	SEQ ID NO:55	SEQ ID NO:56	Fig. 28	PTA-213
rat CARD-6	SEQ ID NO:51	SEQ ID NO:52	SEQ ID NO:53	Fig. 25	PTA-211

A nucleotide sequence encoding a human CARD-3 protein is shown in Figure 1 (SEQ ID NO:1; SEQ ID NO:3 includes the open reading frame only).

A predicted amino acid sequence of CARD-3 protein is also shown in Figure 2

10 (SEQ ID NO:2).

CARD-4 has at least two forms, a long form, CARD-4L, and a short form, CARD-4S, as well as two or more splice variants. A nucleotide sequence encoding a human CARD-4L protein is shown in Figure 3 (SEQ ID NO:7; SEQ ID NO:9 includes the open reading frame only). A predicted amino acid

15 sequence of CARD-4L protein is also shown in Figure 4 (SEQ ID NO:8). A

nucleotide sequence encoding a human CARD-4S protein is shown in Figure 5 (SEQ ID NO:25; SEQ ID NO:27 includes the open reading frame only). A

predicted amino acid sequence of CARD-4S protein is shown in Figure 6 (SEQ ID NO:26). Two additional splice variants of human CARD-4 are provided in Figures 10 and 11 (human CARD-4Y) and Figures 12 and 13 (human CARD-4Z) (predicted amino acid sequences: SEQ ID NO:39 and SEQ ID NO:41 and nucleic acid sequences: SEQ ID NO:38 and SEQ ID NO:40). These two splice variants are predicted to contain 249 and 164 amino acids, respectively. An alignment of human CARD-4Y, human CARD-4Z and human CARD-4L is shown in Figure 14.

In addition to the human CARD-4 proteins, a full length nucleotide sequence of the murine ortholog of human CARD-4L is provided in Figure 15 (SEQ ID NO:42). An alignment of murine CARD-4L with human CARD-4L is shown in Figure 17.

A nucleotide sequence encoding a murine CARD-5 protein is shown in Figure 19 (SEQ ID NO:60; SEQ ID NO:62 includes the open reading frame only). A predicted amino acid sequence of murine CARD-5 protein is also shown in Figure 19 (SEQ ID NO:61).

A nucleotide sequence encoding a human CARD-5 protein is shown in Figure 21 (SEQ ID NO:48; SEQ ID NO:50 includes the open reading frame only). A predicted amino acid sequence of human CARD-5 protein is also shown in Figure 21 (SEQ ID NO:49).

A nucleotide sequence encoding a rat CARD-6 protein is shown in Figure 25 (SEQ ID NO:51; SEQ ID NO:53 includes the open reading frame only). A predicted amino acid sequence of rat CARD-6 protein is also shown in Figure 25 (SEQ ID NO:52).

The human CARD-3 cDNA of Figure 1 (SEQ ID NO:1), which is approximately 1931 nucleotides long including untranslated regions, encodes a protein having a molecular weight of approximately 61 kDa (excluding post-translational modifications).

A plasmid containing a cDNA encoding human CARD-3 (pXE17A) was deposited with the American Type Culture Collection (ATCC), Manassas.

VA on May 14, 1998, and assigned Accession Number 203037. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. § 112.

The human CARD-4L cDNA of Figure 3 (SEQ ID NO:7), which is approximately 3382 nucleotides long including untranslated regions, encodes a protein having a molecular weight of approximately 108 kDa (excluding post-translational modifications).

10 A plasmid containing a cDNA encoding human CARD-4L (pC4L1) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on July 7, 1998, and assigned Accession Number 203035. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. § 112.

The human CARD-4S cDNA of Figure 5 (SEQ ID NO:25), which is 3082 nucleotides long including untranslated regions.

A plasmid containing a cDNA encoding human CARD-4S (pDB33E) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on May 14, 1998, and assigned Accession Number 203036. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. § 112.

The human CARD-5 cDNA of Figure 21 (SEQ ID NO:48), which is approximately 740 nucleotides long including untranslated regions, encodes a protein having a molecular weight of approximately 21.6 kD.

A plasmid containing a cDNA encoding human CARD-5 (EpHC5) was deposited with the American Type Culture Collection (ATCC).

Manassas, VA on June 11, 1999, and assigned Accession Number PTA-213. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The murine CARD-5 cDNA of Figure 19 (SEQ ID NO:60), which is approximately 778 nucleotides long, including untranslated regions, encodes a protein having a molecular weight of approximately 21.5 kD.

10 A plasmid containing a cDNA encoding murine CARD-5 (EpMC5) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on June 11, 1999, and assigned Accession Number PTA-212. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent  
15 Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human CARD-6 cDNA of Figure 28 (SEQ ID NO:54), which is approximately 4244 nucleotides long encodes a protein having a molecular weight of approximately 116.5 kD (excluding post-translational modifications).

20 A plasmid containing a cDNA encoding an amino terminal portion of human CARD-6 (EpHC6e), a plasmid encoding a carboxy terminal portion of human CARD-6 (EpHC6c), and a plasmid containing cDNA encoding human CARD-6 (EpHC6) were deposited with the American Type Culture Collection (ATCC), Manassas, VA on June 11, 1999, and assigned Accession Number PTA-  
25 213. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

30

The rat CARD-6 cDNA of Figure 25 (SEQ ID NO:51), which is approximately 5252 nucleotides long including untranslated regions, encodes a protein having a molecular weight of approximately 100.7 kD.

A plasmid containing a cDNA encoding rat CARD-6 (EpMR5) was deposited with the American Type Culture Collection (ATCC), Manassas, VA. on June 10, 1999, and assigned Accession Number PTA-211. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. 112.

A region of human CARD-4L protein (SEQ ID NO:8), the CARD domain (SEQ ID NO:10), bears some similarity to a CARD domain of CARD-3 (SEQ ID NO:6), ARC-CARD (SEQ ID NO:31), cIAP1-CARD (SEQ ID NO:32), and cIAP2-CARD (SEQ ID NO:33). This comparison is depicted in Figure 7.

A region, the CARD domain (SEQ ID NO:58), of human CARD-5 protein (SEQ ID NO:48) and a region, the CARD domain (SEQ ID NO:57), of murine CARD-5 protein (SEQ ID NO:61) bear some similarity to the CARD of RAIDD (SEQ ID NO:70). This comparison is depicted in Figure 27.

Each of CARD-3, CARD-4, CARD-5, and CARD-6 are members of a family of molecules (the "CARD-3 family", the "CARD-4 family", the "CARD-5 family", and the "CARD-6 family" respectively) having certain conserved structural and functional features. The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Such family members can be naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin and a homologue of that protein of murine origin, as well as a second, distinct protein of human origin and a murine homologue of that protein. Members of a family may also have common functional characteristics.

In one embodiment, a CARD-3, CARD-4, CARD-5, or CARD-6 protein includes a CARD domain having at least about 65%, preferably at least about 75%, and more preferably about 85%, 95%, or 98% amino acid sequence identity to the CARD domain of SEQ ID NO:6 or the CARD domain of SEQ ID NO:10 or the CARD domain of SEQ ID NO:28, the CARD domain of SEQ ID NO:57, the CARD domain of SEQ ID NO:58, the CARD domain of SEQ ID NO:59, or the CARD domain of SEQ ID NO:64.

Preferred CARD-3, CARD-4, CARD-5, or CARD-6 polypeptides of the present invention have an amino acid sequence sufficiently identical to the CARD domain amino acid sequence of SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, and SEQ ID NO:64, respectively.

The CARD-3 polypeptide also has an amino acid sequence sufficiently identical to the kinase domain sequence of SEQ ID NO:4, and an amino acid sequence that is sufficiently identical to the linker domain of SEQ ID NO:5. The CARD-4L polypeptide has an amino acid sequence sufficiently identical to the nucleotide binding domain of SEQ ID NO:11, an amino acid sequence sufficiently identical to the Walker Box "A" of SEQ ID NO:12 or Walker Box "B" of SEQ ID NO:13, an amino acid sequence sufficiently identical to the kinase 1a subdomain of SEQ ID NO:46, an amino acid sequence sufficiently identical to the kinase 2 subdomain of SEQ ID NO:47, or an amino acid sequence sufficiently identical to the kinase 3a subdomain of SEQ ID NO:14, or an amino acid sequence sufficiently identical to the Leucine-rich repeats of SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, and SEQ ID NO:24.

As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have

a common structural domain and/or common functional activity. For example, amino acid or nucleotide sequences which contain a common structural domain having about 65% identity, preferably 75% identity, more preferably 85%, 95%, or 98% identity are defined herein as sufficiently identical.

- 5 As used interchangeably herein a "CARD-3, CARD-4, CARD-5, or CARD-6 activity", "biological activity of CARD-3, CARD-4, CARD-5, or CARD-6" or "functional activity of CARD-3, CARD-4, CARD-5, or CARD-6", refers to an activity exerted by a CARD-3, CARD-4, CARD-5, or CARD-6 protein, polypeptide or nucleic acid molecule on a CARD-3, CARD-4, CARD-5, or CARD-6 responsive cell as determined in vivo, or in vitro, according to standard techniques. A CARD-3, CARD-4, CARD-5, or CARD-6 activity can be a direct activity, such as an association with or an enzymatic activity on a second protein or an indirect activity, such as a cellular signaling activity mediated by interaction of the CARD-3, CARD-4, CARD-5, or CARD-6 protein with a second protein. In one embodiment, a CARD-3, CARD-4, CARD-5, or CARD-6 activity includes at least one or more of the following activities: (i) the ability to interact with proteins in an apoptotic signaling pathway; (ii) the ability to interact with a CARD-3, CARD-4, CARD-5, or CARD-6 ligand; or (iii) the ability to interact with an intracellular target protein; (iv) the ability to interact, directly or indirectly with one or more with caspases; (v) the ability to modulate the activity of a caspase, e.g., caspase-9; (vi) the ability to modulate the activity of NF- $\kappa$ B; (vii) the ability to modulate Apaf-1; (viii) the ability to modulate a Bcl-2 family member; (ix) the ability to modulate a neurotrophin receptor, e.g., P75<sup>NTR</sup>; (x) the ability to modulate the activity of a stress activated kinase (e.g., JNK/p38); and (xi) the ability to modulate phosphorylation of CHOP (GADD 153). For example, in Example 4, CARD-3-containing proteins were shown to associate with CARD-4-containing proteins. In example 9, CARD-4 proteins were shown to induce NF- $\kappa$ B-mediated transcription. In example 10, CARD-3 and CARD-4 were shown to enhance caspase-9 activity.

30



CARD-4 and CARD-6 have Apaf-1-like sequences and may bind to one or more members of the Bcl-2 family (e.g., Bcl-2, Boo, or Diva). CARD-3 and CARD-5 may also bind to one or more members of the Bcl-2 family.

5 CARD-3, CARD-4, CARD-5, and CARD-6 may modulate apoptosis by influencing the activity of a Bcl-2 family member, which modulation, in turn, modulates activity of Apaf-1 or other factors. CARD-3, CARD-4, CARD-5, and CARD-6 nucleic acid and polypeptides as well as modulators of activity of expression of CARD-3, CARD-4, CARD-5, or CARD-6 can be used to modulate an Apaf-1 signaling pathway.

10 CARD-3 and CARD-4 may bind to a neurotrophin receptor (e.g., p75<sup>NTR</sup>). CARD-3 and CARD-4 may modulate the activity of a neurotrophin receptor and thus modulate apoptosis of neuronal cells. Accordingly, CARD-3 and CARD-4 nucleic acids and polypeptides as well as modulators of CARD-3 or CARD-4 activity or expression can be used to modulate apoptosis of neurons  
15 (e.g., for treatment of neurological disorders, particularly neurodegenerative disorders).

Accordingly, another embodiment of the invention features isolated CARD-3, CARD-4, CARD-5, or CARD-6 proteins and polypeptides having a CARD-3, CARD-4, CARD-5, or CARD-6 activity.

20 Various aspects of the invention are described in further detail in the following subsections.

#### I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules  
25 that encode CARD-3, CARD-4, CARD-5, or CARD-6 proteins or biologically active portions thereof, as well as nucleic acid molecules sufficient for use as hybridization probes to identify CARD-3, CARD-4, CARD-5, or CARD-6-encoding nucleic acids (e.g., CARD-3, CARD-4, CARD-5, or CARD-6 mRNA) and fragments for use as PCR primers for the amplification or mutation of  
30 CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules. As used

herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded

5 DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences (preferably protein encoding sequences) which naturally flank the nucleic acid (i.e.,  
10 sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in  
15 genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

20 A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the  
25 cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213, or a complement of any of these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid  
30 sequences of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ

ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037 the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of PTA-213, as a hybridization probe, CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook et al., eds., Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid of the invention can be amplified using cDNA, mRNA or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to CARD-3, CARD-4, CARD-5, or CARD-6 nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, or the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of PTA-212, or the cDNA of PTA-213, or a portion thereof. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, the nucleic acid molecule of the invention can comprise

only a portion of a nucleic acid sequence encoding CARD-3, CARD-4, CARD-5, or CARD-6, for example, a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of CARD-3, CARD-4, CARD-5, or CARD-6. The nucleotide sequence determined from the cloning of the human  
5 CARD-3, CARD-4, CARD-5, or CARD-6, and the partial murine CARD-4 gene allows for the generation of probes and primers designed for use in identifying and/or cloning CARD-3, CARD-4, CARD-5, or CARD-6 homologues in other cell types, e.g., from other tissues, as well as CARD-3, CARD-4, CARD-5, or CARD-6 homologues and orthologs from other mammals. The probe/primer  
10 typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 or 400 consecutive nucleotides of the sense or anti-sense sequence of SEQ ID NO:1, SEQ ID NO:3,  
15 SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, or the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of PTA-212, or  
20 the cDNA of PTA-213, or of a naturally occurring mutant of one of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC  
25 203035, the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of PTA-212, or the cDNA of PTA-213.

Probes based on the CARD-3, CARD-4, CARD-5, or CARD-6 nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or similar proteins. The probe comprises a label group  
30 attached thereto, e.g., a radioisotope, a fluorescent compound, an enzyme, or an

enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying allelic variants and orthologs of the CARD-3, CARD-4, CARD-5, or CARD-6 proteins of the present invention, identifying cells or tissue which mis-express a CARD-3, CARD-4, CARD-5, or CARD-6 protein, such as by

5 measuring a level of a CARD-3, CARD-4, CARD-5, or CARD-6-encoding nucleic acid in a sample of cells from a subject, e.g., detecting CARD-3, CARD-4, CARD-5, or CARD-6 mRNA levels or determining whether a genomic CARD-3, CARD-4, CARD-5, or CARD-6 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion" of

10 CARD-3, CARD-4, CARD-5, or CARD-6 can be prepared by isolating a portion of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA

15 of ATCC 203035, the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of PTA-212, or the cDNA of PTA-213, which encodes a polypeptide having a CARD-3, CARD-4, CARD-5, or CARD-6 biological activity, expressing the encoded portion of CARD-3, CARD-4, CARD-5, or CARD-6 protein (e.g., by recombinant expression in vitro) and assessing the activity of the

20 encoded portion of CARD-3, CARD-4, CARD-5, or CARD-6. For example, a nucleic acid fragment encoding a biologically active portion of CARD-3, CARD-4, CARD-5, or CARD-6 includes a CARD domain, e.g., SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:28, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, or SEQ ID NO:62.

25 The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the

30 cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC

203036, or the cDNA of ATCC PTA-211, the cDNA of PTA-212, or the cDNA of PTA-213, due to degeneracy of the genetic code and thus encode the same CARD-3, CARD-4, CARD-5, or CARD-6 protein as that encoded by the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213.

In addition to the CARD-3, CARD-4, CARD-5, or CARD-6 nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of CARD-3, CARD-4, CARD-5, or CARD-6 may exist within a population (e.g., the human population). Such genetic polymorphism in the CARD-3, CARD-4, CARD-5, or CARD-6 gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a CARD-3, CARD-4, CARD-5, or CARD-6 protein, preferably a mammalian CARD-3, CARD-4, CARD-5, or CARD-6 protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the CARD-3, CARD-4, CARD-5, or CARD-6 gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in CARD-3, CARD-4, CARD-5, or CARD-6 that are the result of natural allelic variation and that do not alter the

functional activity of CARD-3, CARD-4, CARD-5, or CARD-6 are intended to be within the scope of the invention. Thus, e.g., 1%, 2%, 3%, 4%, or 5% of the amino acids in CARD-3, CARD-4, CARD-5, or CARD-6 are replaced by another amino acid, preferably the amino acids are replaced by conservative substitutions.

5           Moreover, nucleic acid molecules encoding CARD-3, CARD-4, CARD-5, or CARD-6 proteins from other species (CARD-3, CARD-4, CARD-5, or CARD-6 orthologs/homologues), which have a nucleotide sequence which differs from that of a CARD-3, CARD-4, CARD-5, or CARD-6 disclosed herein, are intended to be within the scope of the invention.

10           For example, Example 5 describes the murine CARD-4 ortholog and Example 14 describes the murine CARD-5 ortholog. Nucleic acid molecules corresponding to natural allelic variants and homologues of the CARD-3, CARD-4, CARD-5, or CARD-6 cDNA of the invention can be isolated based on their similarity to the nucleic acids disclosed herein using the human or murine  
15 cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

          In general, an allelic variant of a gene will be readily identifiable as mapping to the same chromosomal location as said gene. For example, in Example 6, the chromosomal location of the human CARD-4 gene is discovered  
20 to be chromosome 7 close to the SHGC-31928 genetic marker. Allelic variants of human CARD-4 will be readily identifiable as mapping to the human CARD-4 locus on chromosome 7 near genetic marker SHGC-31928.

          Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (325, 350, 375, 400, 425, 450, 500, 550,  
25 600, 650, 700, 800, 900, 1000, 1300, 1600 or 1931) nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence, preferably the coding sequence, of SEQ ID NO:1, SEQ ID NO:3, or the cDNA of ATCC 203037. In yet another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (325, 350, 375, 400, 425,  
30 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1300, 1640, 1900, 2200, 2500,

2800, 3100, or 3382) nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence, preferably the coding sequence, of SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, the cDNA of ATCC 203035, or  
5 the cDNA of ATCC 203036. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1300, 1640, 1900, 2200, 2500, 2800, 3100, 3300, 3600, 3900, 4200 or 4209) nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the  
10 nucleotide sequence, preferably the coding sequence, of SEQ ID NO:42.

In yet another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (350, 400, 450, 500, 550, 600, 650, 700, or 740) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:48 or SEQ ID  
15 NO:50.

In yet another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (350, 400, 450, 500, 550, 600, 650, 700, or 761) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:60 or SEQ ID  
20 NO:62.

In yet another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (350, 400, 450, 500, 550, 600, 650, 700, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5200, or 5252) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule consisting of  
25 the nucleotide sequence of SEQ ID NO:51 or SEQ ID NO:53.

In yet another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (350, 400, 450, 500, 550, 600, 650, 700, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, or 5000) nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule consisting of the  
30 nucleotide sequence of SEQ ID NO:54 or SEQ ID NO:56.



As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. An non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C (e.g., 50°C or 60°C or 65°C). Preferably, the isolated nucleic acid molecule of the invention that hybridizes under stringent conditions corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In addition to naturally-occurring allelic variants of the CARD-3, CARD-4, CARD-5, or CARD-6 sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213, thereby leading to changes in the amino acid sequence of the encoded protein without altering the functional ability of the protein. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of CARD-3, CARD-4L/S, CARD-4 splice variant, murine CARD-4 protein, human CARD-5 protein, murine CARD-5 protein, or rat CARD-6 protein without

altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the CARD-3, CARD-4L/S, CARD-4 splice variant, CARD-4, CARD-5, or CARD-6 proteins of various species are predicted to be particularly  
5 unamenable to alteration.

For example, preferred CARD-3, CARD-4, CARD-5, and CARD-6 proteins of the present invention, contain at least one CARD domain. Additionally, a CARD-3 protein also contains at least one kinase domain or at least one linker domain. A CARD domain contains at least one nucleotide  
10 binding domain or Leucine-rich repeats. Such conserved domains are less likely to be amenable to mutation. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved among CARD-3, CARD-4, CARD-5, or CARD-6 of various species) may not be essential for activity and thus are likely to be amenable to alteration.

15 Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding CARD-3, CARD-4, CARD-5, or CARD-6 proteins that contain changes in amino acid residues that are not essential for activity. Such CARD-3, CARD-4, CARD-5, or CARD-6 proteins differ in amino acid sequence from SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:25, SEQ ID NO:39, SEQ ID  
20 NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61, and yet retain biological activity. In one embodiment, the isolated nucleic acid molecule includes a nucleotide sequence encoding a protein that includes an amino acid sequence that is at least about 45% identical, 65%, 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:2,  
25 SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61.

An isolated nucleic acid molecule encoding a CARD-3, CARD-4, CARD-5, or CARD-6 protein having a sequence which differs from that of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID:25, SEQ ID  
30 NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ

ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, the cDNA of ATCC 203037, the cDNA of ATCC 203035, the cDNA of ATCC 203036, the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213, can be created by

5 introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of CARD-3 (SEQ ID NO:1, SEQ ID NO:3, the cDNA of ATCC 203037) or CARD-4L (SEQ ID NO:7, SEQ ID NO:9, the cDNA of ATCC 203035), or CARD-4S (SEQ ID NO:25, SEQ ID NO:27, the cDNA of ATCC 203036), or human CARD-4 splice variants (SEQ ID NO:38, SEQ ID NO:40, or

10 murine CARD-4 (SEQ ID NO:42), or murine CARD-5 (SEQ ID NO:60, SEQ ID NO:62, the cDNA of PTA-211), or human CARD-5 (SEQ ID NO:48, SEQ ID NO:50, the cDNA of ATCC PTA-213), rat CARD-6 (SEQ ID NO:51, SEQ ID NO:53, the cDNA of ATCC PTA-211), or human CARD-6 (SEQ ID NO:54, SEQ ID NO:56, the cDNA of ATCC PTA-213) such that one or more amino acid

15 substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. Thus, for example, 1%, 2%, 3%, 5%, or 10% of the amino acids can be

20 replaced by conservative substitution. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g.,

25 aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus,

30 a predicted nonessential amino acid residue in CARD-3, CARD-4, CARD-5, or

CARD-6 is preferably replaced with another amino acid residue from the same side chain family. Alternatively, mutations can be introduced randomly along all or part of a CARD-3, CARD-4, CARD-5, or CARD-6 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for CARD-  
5 3, CARD-4, CARD-5, or CARD-6 biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In an embodiment, a mutant CARD-3, CARD-4, CARD-5, or CARD-6 protein can be assayed for: (1) the ability to form protein:protein interactions  
10 with proteins in the apoptotic signaling pathway; (2) the ability to bind a CARD-3, CARD-4, CARD-5, or CARD-6 ligand; or (3) the ability to bind to an intracellular target protein. For example, (1) in Example 7, a two-hybrid screening assay for the physical interaction of CARD-3 and CARD-4 is shown, (2) in Example 8, a two-hybrid system assay for the interaction between CARD-4  
15 and its ligand hNUDC is described, and (3) in Example 12, a coimmunoprecipitation assay for the interaction of CARD-3 with its ligand CARD-4 is shown. In yet another embodiment, a mutant CARD-3, CARD-4, CARD-5, or CARD-6 protein can be assayed for the ability to modulate cellular proliferation, cellular differentiation, or cellular death. For example, in Example  
20 10, assays for the regulation of cellular death (apoptosis) by CARD-3 or CARD-4 are described. In yet another embodiment, a mutant CARD-3 or CARD-4 protein can be assayed for regulation of a cellular signal transduction pathway. For example, in Example 9, an assay for the regulation by CARD-4 of the NF- $\kappa$ B pathway is described.

25 The present invention encompasses antisense nucleic acid molecules, i.e., molecules which are complementary to a sense nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic  
30 acid can be complementary to an entire CARD-3, CARD-4, CARD-5, or CARD-

6 coding strand, or to only a portion thereof, e.g., all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can be antisense to a noncoding region of the coding strand of a nucleotide sequence encoding CARD-3, CARD-4, CARD-5, or CARD-6. The noncoding regions ("5'  
5 and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids. Given the coding strand sequences encoding CARD-3, CARD-4, CARD-5, and CARD-6 disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can  
10 be complementary to the entire coding region of CARD-3, CARD-4, CARD-5, or CARD-6L/S mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of CARD-3  
15 mRNA, e.g., an oligonucleotide having the sequence  
CCCTGGTACTTGCCCCCTCCGGTAG (SEQ ID NO:34) or  
CCTGGTACTTGCCCCCTCC (SEQ ID NO:35) or of the CARD-4L mRNA, e.g.,  
TCGTTAAGCCCTTGAAGACAGTG (SEQ ID NO:36) and  
TCGTTAGCCCTTGAAGACCAGTGAGTGTAG (SEQ ID NO:37) or of the  
20 human CARD-5 mRNA, e.g., TAGGACCTCGGTACCCGCGCGCGCG (SEQ ID NO:68) or CGCCGGCCCC TAGGACCTCGGTACC (SEQ ID NO:69). An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using  
25 procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g.,  
30 phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, 5 dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-10 methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-aino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid 15 can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically 20 administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a CARD-3, CARD-4, CARD-5, or CARD-6 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for 25 example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then 30 administered systemically. For example, for systemic administration, antisense

molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered  
5 to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an  $\alpha$ -  
10 anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids. Res. 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res.  
15 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) FEBS Lett. 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a  
20 complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-91)) can be used to catalytically cleave CARD-3, CARD-4, CARD-5, or CARD-6 mRNA transcripts to thereby inhibit translation of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA. A ribozyme having specificity for a CARD-3, CARD-4, CARD-5, or  
25 CARD-6-encoding nucleic acid can be designed based upon the nucleotide sequence of a CARD-3, CARD-4, CARD-5, or CARD-6 cDNA disclosed herein.

For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a CARD-3, CARD-4, CARD-5, or CARD-  
30 6-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et

al. U.S. Patent No. 5,116,742. Alternatively, CARD-3, CARD-4, CARD-5, or CARD-6 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel and Szostak (1993) *Science* 261:1411-1418.

5           The invention also encompasses nucleic acid molecules which form triple helical structures. For example, CARD-3, CARD-4, CARD-5, or CARD-6 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the CARD-3, CARD-4, CARD-5, or CARD-6 (e.g., the CARD-3, CARD-4, CARD-5, or CARD-6 promoter and/or  
10 enhancers) to form triple helical structures that prevent transcription of the CARD-3, CARD-4, CARD-5, or CARD-6 gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

          In embodiments, the nucleic acid molecules of the invention can be  
15 modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) *Bioorganic & Medicinal Chemistry* 4(1):5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to  
20 nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid  
25 phase peptide synthesis protocols as described in Hyrup et al. (1996) *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

          PNAs of CARD-3, CARD-4, CARD-5, or CARD-6 can be used for therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression  
30 by, e.g., inducing transcription or translation arrest or inhibiting replication.



PNAs of CARD-3, CARD-4, CARD-5, or CARD-6 can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup (1996) supra; or as probes or primers for

5 DNA sequence and hybridization (Hyrup (1996) supra; Perry-O'Keefe et al. (1996) Proc. Natl. Acad. Sci. USA 93: 14670-675).

In another embodiment, PNAs of CARD-3, CARD-4, CARD-5, or CARD-6 can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-

10 DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of CARD-3, CARD-4, CARD-5, or CARD-6 can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the

15 PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) supra). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) supra and Finn et al. (1996) Nucleic Acids Research

20 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag et al. (1989) Nucleic Acid Res. 17:5973-88). PNA monomers are then coupled in

25 a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) Nucleic Acids Research 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser et al. (1975) Bioorganic Med. Chem. Lett. 5:1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W0 88/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W0 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, e.g., Krol et al. (1988) *Bio/Techniques* 6:958-976) or intercalating agents (see, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide  
10 may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

## II. Isolated CARD-3, CARD-4, CARD-5, and CARD-6 Proteins and Anti-CARD-3, CARD-4, CARD-5, and CARD-6 Antibodies.

15 One aspect of the invention pertains to isolated CARD-3, CARD-4, CARD-5, and CARD-6 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies. In one embodiment, native CARD-3, CARD-4, CARD-5, or CARD-6 proteins can be isolated from cells or tissue  
20 sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, CARD-3, CARD-4, CARD-5, or CARD-6 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a CARD-3, CARD-4, CARD-5, or CARD-6 protein or polypeptide can be synthesized chemically using standard peptide synthesis  
25 techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the CARD-3, CARD-4, CARD-5, or CARD-6 protein is derived, or substantially free from chemical precursors or  
30 other chemicals when chemically synthesized. The language "substantially free

of cellular material" includes preparations of CARD-3, CARD-4, CARD-5, or CARD-6 protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, CARD-3, CARD-4, CARD-5, or CARD-6 protein that is substantially free of cellular

5 material includes preparations of CARD-3, CARD-4, CARD-5, or CARD-6 protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of non-CARD-3, CARD-4, CARD-5, or CARD-6 protein (also referred to herein as a "contaminating protein"). When the CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof is recombinantly produced, it is also

10 preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When CARD-3, CARD-4, CARD-5, or CARD-6 protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, i.e., it is separated from chemical precursors or other chemicals which

15 are involved in the synthesis of the protein. Accordingly such preparations of CARD-3, CARD-4, CARD-5, or CARD-6 protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or non-CARD-3, CARD-4, CARD-5, or CARD-6 chemicals.

Biologically active portions of a CARD-3, CARD-4, CARD-5, or

20 CARD-6 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the CARD-3, CARD-4, CARD-5, or CARD-6 protein (e.g., the amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61),

25 which include less amino acids than the full length CARD-3, CARD-4, CARD-5, or CARD-6 protein, and exhibit at least one activity of a CARD-3, CARD-4, CARD-5, or CARD-6 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the CARD-3, CARD-4, CARD-5, or CARD-6 protein. A biologically active portion of a CARD-3, CARD-4, CARD-

30 5, or CARD-6 protein can be a polypeptide which is, for example, 10, 25, 50, 100

or more amino acids in length. Preferred biologically active polypeptides include one or more identified CARD-3, CARD-4, CARD-5, or CARD-6 structural domains, e.g., the CARD domain (SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:27, SEQ ID NO:66, SEQ ID NO:67, or SEQ ID NO:68).

5 Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native CARD-3, CARD-4, CARD-5, or CARD-6 protein.

CARD-3, CARD-4, CARD-5, or CARD-6 protein has the amino acid  
10 sequence shown of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61. Other useful CARD-3, CARD-4, CARD-5, or CARD-6 proteins are substantially identical to SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID  
15 NO:49, SEQ ID NO:52, or SEQ ID NO:55, or SEQ ID NO:61, and retain the functional activity of the protein of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61, yet differ in amino acid sequence due to natural allelic variation or mutagenesis. CARD-3 and CARD-4 are  
20 involved in activating caspases in the apoptotic pathway. For example, in Example 10, CARD-4 is shown to enhance caspase 9 activity.

A useful CARD-3, CARD-4, CARD-5, or CARD-6 protein is a protein which includes an amino acid sequence at least about 45%, preferably 55%, 65%, 75%, 85%, 95%, or 99% identical to the amino acid sequence of SEQ  
25 ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61, and retains the functional activity of the CARD-3, CARD-4, CARD-5, or CARD-6 proteins of SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID  
30 NO:55, or SEQ ID NO:61.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence).

- 5 The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of  
10 identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions x 100).

- The determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two  
15 sequences is the algorithm of Karlin and Altschul (1990) Proc. Nat'l Acad. Sci. USA 87:2264-2268, modified as in Karlin and Altschul (1993) Proc. Nat'l Acad. Sci. USA 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) J. Mol. Biol. 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score  
20 = 100, wordlength = 12 to obtain nucleotide sequences similar or homologous to CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules of the invention. For example, Example 5 describes the use of the TBLASTN program to query a database of sequences of full length and partial cDNA sequences with the human CARD-4 polypeptide sequence leading to the discovery of murine  
25 CARD-4 and Example 4 describes the use of BLASTN to query a proprietary EST database with the 5' untranslated sequence of CARD-4 leading to the discovery of two human CARD-4 splice variants. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to CARD-3, CARD-4, CARD-5, or CARD-6  
30 protein molecules of the invention. To obtain gapped alignments for comparison

purposes. Gapped BLAST can be utilized as described in Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred.

- 5 non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller. CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a  
10 gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

- The invention also provides CARD-3, CARD-4, CARD-5, or CARD-  
15 6 chimeric or fusion proteins. As used herein, a CARD-3, CARD-4, CARD-5, or CARD-6 "chimeric protein" or "fusion protein" comprises a CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide operatively linked to a non-CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide. A "CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide" refers to a polypeptide having an amino acid sequence  
20 corresponding to all or a portion (preferably a biologically active portion) of a CARD-3, CARD-4, CARD-5, or CARD-6, whereas a "non-CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially identical to the CARD-3, CARD-4, CARD-5, or CARD-6 protein. e.g., a protein which is  
25 different from the CARD-3, CARD-4, CARD-5, or CARD-6 proteins and which is derived from the same or a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide and the non-CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide are fused in-frame to each other. The heterologous  
30 polypeptide can be fused to the N-terminus or C-terminus of the CARD-3.

CARD-4, CARD-5, or CARD-6 polypeptide.

One useful fusion protein is a GST fusion protein in which the CARD-3, CARD-4, CARD-5, or CARD-6 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of  
5 recombinant CARD-3, CARD-4, CARD-5, or CARD-6. In another embodiment, the fusion protein contains a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of CARD-3, CARD-4, CARD-5, or CARD-6 can be increased through use of a heterologous signal sequence. For example, the gp67 secretory sequence of the baculovirus  
10 envelope protein can be used as a heterologous signal sequence (Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal  
15 sequences include the phoA secretory signal (Molecular cloning, Sambrook et al. second edition, Cold spring harbor laboratory press, 1989) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is a CARD-3, CARD-4, CARD-5, or CARD-6-immunoglobulin fusion protein in which all or part of  
20 CARD-3, CARD-4, CARD-5, or CARD-6 is fused to sequences derived from a member of the immunoglobulin protein family. The CARD-3, CARD-4, CARD-5, or CARD-6-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a CARD-3, CARD-4, CARD-5, or CARD-6 ligand  
25 and a CARD-3, CARD-4, CARD-5, or CARD-6 protein on the surface of a cell, to thereby suppress CARD-3, CARD-4, CARD-5, or CARD-6-mediated signal transduction in vivo. The CARD-3, CARD-4, CARD-5, or CARD-6-immunoglobulin fusion proteins can be used to affect the bioavailability of a CARD-3, CARD-4, CARD-5, or CARD-6 cognate ligand. Inhibition of the  
30 CARD-3 ligand/CARD-3, CARD-4 ligand/CARD-4, CARD-5 ligand/CARD-5,

or CARD-6 ligand/CARD-6 interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the CARD-3, CARD-4, CARD-5, or CARD-6-immunoglobulin fusion proteins of the invention can be  
5 used as immunogens to produce anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies in a subject, to purify CARD-3, CARD-4, CARD-5, or CARD-6 ligands and in screening assays to identify molecules which inhibit the interaction of CARD-3, CARD-4, CARD-5, or CARD-6 with a CARD-3, CARD-4, CARD-5, or CARD-6 ligand.

10 Preferably, a CARD-3, CARD-4, CARD-5, or CARD-6 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for  
15 ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried  
20 out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, e.g., Current Protocols in Molecular Biology, Ausubel et al. eds., John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion  
25 moiety (e.g., a GST polypeptide). A CARD-3, CARD-4, CARD-5, or CARD-6-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the CARD-3, CARD-4, CARD-5, or CARD-6 protein.

The present invention also pertains to variants of the CARD-3,  
30 CARD-4, CARD-5, or CARD-6 proteins which function as either CARD-3,



CARD-4, CARD-5, or CARD-6 agonists (mimetics) or as CARD-3, CARD-4, CARD-5, or CARD-6 antagonists. Variants of the CARD-3, CARD-4, CARD-5, or CARD-6 protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the CARD-3, CARD-4, CARD-5, or CARD-6 protein.

- 5 An agonist of the CARD-3, CARD-4, CARD-5, or CARD-6 protein can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the CARD-3, CARD-4, CARD-5, or CARD-6 protein. An antagonist of the CARD-3, CARD-4, CARD-5, or CARD-6 protein can inhibit one or more of the activities of the naturally occurring form of the CARD-3,  
10 CARD-4, CARD-5, or CARD-6 protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the CARD-3, CARD-4, CARD-5, or CARD-6 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities  
15 of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the CARD-3, CARD-4, CARD-5, or CARD-6 proteins.

- Variants of the CARD-3, CARD-4, CARD-5, or CARD-6 protein which function as either CARD-3, CARD-4, CARD-5, or CARD-6 agonists  
20 (mimetics) or as CARD-3, CARD-4, CARD-5, or CARD-6 antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants of the CARD-3, CARD-4, CARD-5, or CARD-6 protein for CARD-3, CARD-4, CARD-5, or CARD-6 protein agonist or antagonist activity. In one embodiment, a variegated library of CARD-3, CARD-4, CARD-5, or CARD-6  
25 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of CARD-3, CARD-4, CARD-5, or CARD-6 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential CARD-3, CARD-4, CARD-5, or CARD-6 sequences  
30 is expressible as individual polypeptides, or alternatively, as a set of larger fusion

proteins (e.g., for phage display) containing the set of CARD-3, CARD-4, CARD-5, or CARD-6 sequences therein. There are a variety of methods which can be used to produce libraries of potential CARD-3, CARD-4, CARD-5, or CARD-6 variants from a degenerate oligonucleotide sequence. Chemical  
5 synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential CARD-3, CARD-4, CARD-5, or CARD-6 sequences. Methods for synthesizing degenerate  
10 oligonucleotides are known in the art (see, e.g., Narang (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477).

Useful fragments of CARD-3, CARD-4, CARD-5, and CARD-6, include fragments comprising or consisting of a domain or subdomain described  
15 herein, e.g., a kinase domain or a CARD domain.

In addition, libraries of fragments of the CARD-3, CARD-4, CARD-5, or CARD-6 protein coding sequence can be used to generate a variegated population of CARD-3, CARD-4, CARD-5, or CARD-6 fragments for screening and subsequent selection of variants of a CARD-3, CARD-4, CARD-5, or  
20 CARD-6 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a CARD-3, CARD-4, CARD-5, or CARD-6 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which  
25 can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the CARD-3, CARD-4, CARD-5, or CARD-6  
30 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the

5 combinatorial mutagenesis of CARD-3, CARD-4, CARD-5, or CARD-6 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes

10 under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify CARD-3, CARD-4, CARD-5, or CARD-6 variants (Arkin and Yourvan

15 (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

An isolated CARD-3, CARD-4, CARD-5, or CARD-6 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind CARD-3, CARD-4, CARD-5, or CARD-6 using standard techniques for

20 polyclonal and monoclonal antibody preparation. The full-length CARD-3, CARD-4, CARD-5, or CARD-6 protein can be used or, alternatively, the invention provides antigenic peptide fragments of CARD-3, CARD-4, CARD-5, or CARD-6 for use as immunogens. The antigenic peptide of CARD-3, CARD-4, CARD-5, or CARD-6 comprises at least 8 (preferably 10, 15, 20, or 30) amino

25 acid residues of the amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:8, SEQ ID NO:26, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:49, SEQ ID NO:52, SEQ ID NO:55, or SEQ ID NO:61 or polypeptides including amino acids 128-139 or 287-298 of human CARD-4L and encompasses an epitope of CARD-3, CARD-4, CARD-5, or CARD-6 such that an antibody

30 raised against the peptide forms a specific immune complex with CARD-3.

CARD-4, CARD-5, or CARD-6.

Useful antibodies include antibodies which bind to a domain or subdomain of CARD-3, CARD-4, CARD-5, or CARD-6 described herein (e.g., a kinase domain, a CARD domain, or a leucine-rich domain).

5 Preferred epitopes encompassed by the antigenic peptide are regions of CARD-3, CARD-4, CARD-5, or CARD-6 that are located on the surface of the protein, e.g., hydrophilic regions. Other important criteria include a preference for a terminal sequence, high antigenic index (e.g., as predicted by Jameson-Wolf algorithm), ease of peptide synthesis (e.g., avoidance of prolines);  
10 and high surface probability (e.g., as predicted by the Emini algorithm; Figure 8 and Figure 9).

A CARD-3, CARD-4, CARD-5, or CARD-6 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic  
15 preparation can contain, for example, recombinantly expressed CARD-3, CARD-4, CARD-5, or CARD-6 protein or a chemically synthesized CARD-3, CARD-4, CARD-5, or CARD-6 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an  
20 immunogenic CARD-3, CARD-4, CARD-5, or CARD-6 preparation induces a polyclonal anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody response. For example, polypeptides including amino acids 128-139 or 287-298 of human CARD-4L were conjugated to KLH and the resulting conjugates were used to immunize rabbits and polyclonal antibodies that specifically recognize the two  
25 immunogen peptides were generated.

Accordingly, another aspect of the invention pertains to anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site  
30 which specifically binds an antigen, such as CARD-3, CARD-4, CARD-5, or

CARD-6. A molecule which specifically binds to CARD-3, CARD-4, CARD-5, or CARD-6 is a molecule which binds CARD-3, CARD-4, CARD-5, or CARD-6, but does not substantially bind other molecules in a sample, e.g., a biological sample, which naturally contains CARD-3, CARD-4, CARD-5, or CARD-6.

- 5 Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')<sub>2</sub> fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind CARD-3, CARD-4, CARD-5, or CARD-6. The term "monoclonal antibody" or "monoclonal antibody composition", as used
- 10 herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of CARD-3, CARD-4, CARD-5, or CARD-6. A monoclonal antibody composition thus typically displays a single binding affinity for a particular CARD-3, CARD-4, CARD-5, or CARD-6 protein with which it immunoreacts.
- 15 Polyclonal anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies can be prepared as described above by immunizing a suitable subject with a CARD-3, CARD-4, CARD-5, or CARD-6 immunogen. The anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked
- 20 immunosorbent assay (ELISA) using immobilized CARD-3, CARD-4, CARD-5, or CARD-6. If desired, the antibody molecules directed against CARD-3, CARD-4, CARD-5, or CARD-6 can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after
- 25 immunization, e.g., when the anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor et al. (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole et al. (1985),
- 30

Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing various antibodies monoclonal antibody hybridomas is well known (see generally Current Protocols in Immunology (1994) Coligan et al. (eds.) John Wiley & Sons, Inc., New York, NY). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a CARD-3, CARD-4, CARD-5, or CARD-6 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds CARD-3, CARD-4, CARD-5, or CARD-6.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-CARD-3, CARD-4, CARD-5, or CARD-6 monoclonal antibody (see, e.g., Current Protocols in Immunology, supra; Galfre et al. (1977) Nature 266:55052; R.H. Kenneth, in Monoclonal Antibodies: A New Dimension In Biological Analyses, Plenum Publishing Corp., New York, New York (1980); and Lerner (1981) Yale J. Biol. Med., 54:387-402). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line, e.g., a myeloma cell line that is sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused

myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind CARD-3, CARD-4, CARD-5, or CARD-6, e.g., using a  
5 standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with CARD-3,  
10 CARD-4, CARD-5, or CARD-6 to thereby isolate immunoglobulin library members that bind CARD-3, CARD-4, CARD-5, or CARD-6. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia Recombinant Phage Antibody System, Catalog No. 27-9400-01; and the Stratagene SurfZAP Phage Display Kit, Catalog No. 240612).  
15 Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO  
20 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffiths et al. (1993) *EMBO J.* 12:725-734.

Additionally, recombinant anti-CARD-3, CARD-4, CARD-5, or  
25 CARD-6 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods  
30 described in PCT Publication No. WO 87/02671; European Patent Application

- 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al. (1988) Science 240:1041-1043; Liu et al. (1987) Proc. Natl. Acad. Sci. USA 84:3439-3443; Liu et al. (1987) J. Immunol. 139:3521-3526; Sun et al. (1987) Proc. Natl. Acad. Sci. USA 84:214-218; Nishimura et al. (1987) Canc. Res. 47:999-1005; Wood et al. (1985) Nature 314:446-449; and Shaw et al. (1988) J. Natl. Cancer Inst. 80:1553-1559; Morrison, (1985) Science 229:1202-1207; Oi et al. (1986) Bio/Techniques 4:214; U.S. Patent 5,225,539; Jones et al. (1986) Nature 321:552-525; Verhoeyan et al. (1988) Science 239:1534; and Beidler et al. (1988) J. Immunol. 141:4053-4060.

An anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody (e.g., monoclonal antibody) can be used to isolate CARD-3, CARD-4, CARD-5, or CARD-6 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody can facilitate the purification of natural CARD-3, CARD-4, CARD-5, or CARD-6 from cells and of recombinantly produced CARD-3, CARD-4, CARD-5, or CARD-6 expressed in host cells. Moreover, an anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody can be used to detect CARD-3, CARD-4, CARD-5, or CARD-6 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the CARD-3, CARD-4, CARD-5, or CARD-6 protein. Anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase,  $\beta$ -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials



include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive  
5 material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ .

### III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding CARD-3, CARD-4,  
10 CARD-5, or CARD-6 (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA  
15 segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are  
20 replicated along with the host genome. Moreover, certain vectors, expression vectors, are capable of directing the expression of genes to which they are operatively linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors  
25 (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one  
30 or more regulatory sequences, selected on the basis of the host cells to be used for

expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide  
5 sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; Gene Expression Technology: Methods in Enzymology  
10 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can  
15 depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., CARD-3, CARD-4, CARD-5, or CARD-6 proteins, mutant forms of CARD-3,  
20 CARD-4, CARD-5, or CARD-6, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of CARD-3, CARD-4, CARD-5, or CARD-6 in prokaryotic or eukaryotic cells, e.g., bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host  
25 cells are discussed further in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the  
30

expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) Gene 69:301-315) and pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident  $\lambda$  prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence

of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis

5 techniques.

In another embodiment, the CARD-3, CARD-4, CARD-5, or CARD-6 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al. (1987) *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88  
10 (Schultz et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), pGBT9 (Clontech, Palo Alto, CA), pGAD10 (Clontech, Palo Alto, CA), pYADE4 and pYGAE2 and pYPGE2 (Brunelli and Pall, (1993) *Yeast* 9:1299-1308), pYPGE15 (Brunelli and Pall, (1993) *Yeast* 9:1309-1318), pACTII (Dr. S.E. Elledge, Baylor College of Medicine), and picZ (InVitrogen Corp, San  
15 Diego, CA). For example, in Example 7 the expression of a fusion protein comprising amino acids 1-145 of human CARD-4L fused to the DNA-binding domain of *S. cerevisiae* transcription factor GAL4 from the yeast expression vector pGBT9 is described. In another example, Example 8 describes the expression of a fusion protein comprising amino acids 406-953 of human CARD-  
20 4L fused to the DNA-binding domain of *S. cerevisiae* transcription factor GAL4 from the yeast expression vector pGBT9. In yet another example, Example 7 describes the expression of a fusion protein comprising CARD-3 fused to the transcriptional activation domain of *S. cerevisiae* transcription factor GAL4 from the yeast expression vector pACTII.

25 Alternatively, CARD-3, CARD-4, CARD-5, or CARD-6 can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

30 In yet another embodiment, a nucleic acid of the invention is

expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) Nature 329:840), pCI (Promega), and pMT2PC (Kaufman et al. (1987) EMBO J. 6:187-195). When used in mammalian cells, the expression vector's control functions are

5 often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook et al. (supra). For example, Example 9, Example 10, and Example 12 describe the expression of human

10 CARD-4 or fragments thereof, CARD-3, or both from the mammalian expression vector pCI.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express

15 the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) Genes Dev. 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) Adv. Immunol. 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) EMBO J.

20 8:729-733) and immunoglobulins (Banerji et al. (1983) Cell 33:729-740; Queen and Baltimore (1983) Cell 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) Proc. Natl. Acad. Sci. USA 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) Science 230:912-916), and mammary gland-specific promoters (e.g., milk whey

25 promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) Science 249:374-379) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman (1989) Genes Dev. 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of  
5 the DNA molecule) of an RNA molecule which is antisense to CARD-3, CARD-4, CARD-5, or CARD-6 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be  
10 chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is  
15 introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al. (Reviews - Trends in Genetics, Vol. 1(1) 1986).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention or isolated nucleic acid molecule of the invention has been introduced. The terms "host cell" and "recombinant  
20 host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the  
25 term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, CARD-3, CARD-4, CARD-5, or CARD-6 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to  
30 those skilled in the art. For example, in Example 7 a *Saccharomyces cerevisiae*

host cell for recombinant CARD-4 and CARD-3 expression is described, and in Examples 9, 10, and 12, a 293T host cells for expression of CARD-4 or fragments thereof or CARD-3 are described.

Vector DNA or an isolated nucleic acid molecule of the invention can  
5 be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated  
10 transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (supra), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a  
15 small fraction of cells may integrate the foreign DNA into their genome. In some cases vector DNA is retained by the host cell. In other cases the host cell does not retain vector DNA and retains only an isolated nucleic acid molecule of the invention carried by the vector. In some cases, an isolated nucleic acid molecule of the invention is used to transform a cell without the use of a vector.

20 In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a  
25 host cell on the same vector as that encoding CARD-3, CARD-4, CARD-5, or CARD-6 or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

30

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) a CARD-3, CARD-4, CARD-5, or CARD-6 protein. Accordingly, the invention further provides methods for producing CARD-3, CARD-4, CARD-5, or CARD-6 protein using  
5 the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector or isolated nucleic acid molecule encoding CARD-3, CARD-4, CARD-5, or CARD-6 has been introduced) in a suitable medium such that CARD-3, CARD-4, CARD-5, or CARD-6 protein is produced. In another embodiment, the  
10 method further comprises isolating CARD-3, CARD-4, CARD-5, or CARD-6 from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which CARD-3, CARD-4,  
15 CARD-5, or CARD-6-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous CARD-3, CARD-4, CARD-5, or CARD-6 sequences have been introduced into their genome or homologous recombinant animals in which endogenous CARD-3, CARD-4, CARD-5, or CARD-6 sequences have been altered. Such animals  
20 are useful for studying the function and/or activity of CARD-3, CARD-4, CARD-5, or CARD-6 and for identifying and/or evaluating modulators of CARD-3, CARD-4, CARD-5, or CARD-6 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a  
25 transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or  
30 tissues of the transgenic animal. As used herein, an "homologous recombinant



animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous CARD-3, CARD-4, CARD-5, or CARD-6 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic  
5 cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing CARD-3, CARD-4, CARD-5, or CARD-6-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The  
10 CARD-3, CARD-4, CARD-5, or CARD-6 cDNA sequence, e.g., that of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, SEQ ID NO:62, or the cDNA of ATCC 203037, or the cDNA of  
15 ATCC 203035, or the cDNA of ATCC 203036, or the cDNA of ATCC PTA-211, the cDNA of ATCC PTA-212, or the cDNA of ATCC PTA-213) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homolog or ortholog of the human CARD-3, CARD-4, CARD-5, or CARD-6 gene, such as a mouse CARD-3, CARD-4, CARD-5, or CARD-6 gene,  
20 can be isolated based on hybridization to the human CARD-3, CARD-4, CARD-5, or CARD-6 cDNA and used as a transgene. For example, the mouse ortholog of CARD-4, Figure 15 and SEQ ID NO:42 can be used to make a transgenic animal using standard methods. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of  
25 the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the CARD-3, CARD-4, CARD-5, or CARD-6 transgene to direct expression of CARD-3, CARD-4, CARD-5, or CARD-6 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are  
30 described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, U.S. Patent

No. 4,873,191 and in Hogan, Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the CARD-3, CARD-4, CARD-5, or  
5 CARD-6 transgene in its genome and/or expression of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding CARD-3, CARD-4, CARD-5, or CARD-6 can further be bred to other transgenic animals  
10 carrying other transgenes.

To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a CARD-3, CARD-4, CARD-5, or CARD-6 gene (e.g., a human or a non-human homolog of the CARD-3, CARD-4, CARD-5, or CARD-6 gene, e.g., a murine CARD-3, CARD-4, CARD-5, or CARD-6  
15 gene) into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the CARD-3, CARD-4, CARD-5, or CARD-6 gene. In an embodiment, the vector is designed such that, upon homologous recombination, the endogenous CARD-3, CARD-4, CARD-5, or CARD-6 gene is functionally disrupted (i.e., no longer encodes a functional  
20 protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous CARD-3, CARD-4, CARD-5, or CARD-6 gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous CARD-3, CARD-4, CARD-5, or  
25 CARD-6 protein). In the homologous recombination vector, the altered portion of the CARD-3, CARD-4, CARD-5, or CARD-6 gene is flanked at its 5' and 3' ends by additional nucleic acid of the CARD-3, CARD-4, CARD-5, or CARD-6 gene to allow for homologous recombination to occur between the exogenous CARD-3, CARD-4, CARD-5, or CARD-6 gene carried by the vector and an endogenous  
30 CARD-3, CARD-4, CARD-5, or CARD-6 gene in an embryonic stem cell. The

additional flanking CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see, e.g., Thomas and Capecchi (1987) Cell 51:503 for  
5 a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced CARD-3, CARD-4, CARD-5, or CARD-6 gene has homologously recombined with the endogenous CARD-3, CARD-4, CARD-5, or CARD-6 gene are selected (see, e.g., Li et al. (1992) Cell 69:915). The selected cells are then  
10 injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see, e.g., Bradley in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the  
15 homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) Current Opinion in Bio/Technology 2:823-829 and in  
20 PCT Publication Nos. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system  
25 of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso et al. (1992) Proc. Natl. Acad. Sci. USA 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) Science 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene,  
30 animals containing transgenes encoding both the Cre recombinase and a selected

protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

5 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut et al. (1997) Nature 385:810-813 and PCT Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter Go phase. The quiescent cell can then  
10 be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the  
15 cell, e.g., the somatic cell, is isolated.

#### IV. Pharmaceutical Compositions

The CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules, CARD-3, CARD-4, CARD-5, or CARD-6 proteins, and anti-CARD-  
20 3, CARD-4, CARD-5, or CARD-6 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is  
25 intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except  
30 compound, use thereof in the compositions is contemplated. Supplementary

active compounds can also be incorporated into the compositions.

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid of the invention. Such methods comprise formulating a  
5 pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or  
10 activity of a polypeptide or nucleic acid of the invention and one or more additional active compounds.

The agent which modulates expression or activity may, for example, be a small molecule. For example, such small molecules include peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides,  
15 polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000  
20 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small  
25 molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention. Exemplary doses include milligram or microgram  
30 amounts of the small molecule per kilogram of subject or sample weight (e.g.,

about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to modulated.  
5 Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for  
10 example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the  
15 route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous,  
20 oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such  
25 as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation  
30 can be enclosed in ampoules, disposable syringes or multiple dose vials made of

glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For  
5 intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL? (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the  
10 contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the  
15 maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as  
20 mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active  
25 compound (e.g., a CARD-3, CARD-4, CARD-5, or CARD-6 protein or anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which  
30 contains a basic dispersion medium and the required other ingredients from those

enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

- 5           Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a
- 10   mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline
- 15   cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring. For administration by
- 20   inhalation, the compounds are delivered in the form of an aerosol spray from a pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

- Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to
- 25   the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are
- 30   formulated into ointments, salves, gels, or creams as generally known in the art.



The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers  
5 that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such  
10 formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those  
15 skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a  
20 predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art  
25 of compounding such an active compound for the treatment of individuals.

For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life  
30 within the human body than other antibodies. Accordingly, lower dosages and

less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) J. Acquired Immune Deficiency Syndromes and Human  
5 Retrovirology 14:193).

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. Patent 5,328,470) or by stereotactic injection (see, e.g., Chen et al. (1994) Proc.  
10 Natl. Acad. Sci. USA 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g. retroviral vectors, the  
15 pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

## 20 V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods:  
a) screening assays; b) detection assays (e.g., chromosomal mapping, tissue  
typing, forensic biology), c) predictive medicine (e.g., diagnostic assays,  
25 prognostic assays, monitoring clinical trials, and pharmacogenomics); and d) methods of treatment (e.g., therapeutic and prophylactic). A CARD-3, CARD-4, CARD-5, or CARD-6 protein interacts with other cellular proteins and can thus be used for (i) regulation of cellular proliferation; (ii) regulation of cellular differentiation; and (iii) regulation of cell survival. The isolated nucleic acid  
30 molecules of the invention can be used to express CARD-3, CARD-4, CARD-5,

or CARD-6 protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect CARD-3, CARD-4, CARD-5, or CARD-6 mRNA (e.g., in a biological sample) or a genetic lesion in a CARD-3, CARD-4, CARD-5, or CARD-6 gene, and to modulate CARD-3, CARD-4, CARD-5, or CARD-6 activity. In addition, the CARD-3, CARD-4, CARD-5, or CARD-6 proteins can be used to screen drugs or compounds which modulate the CARD-3, CARD-4, CARD-5, or CARD-6 activity or expression as well as to treat disorders characterized by insufficient or excessive production of CARD-3, CARD-4, CARD-5, or CARD-6 protein or production of CARD-3, CARD-4, CARD-5, or CARD-6 protein forms which have decreased or aberrant activity compared to CARD-3, CARD-4, CARD-5, or CARD-6 wild type protein. In addition, the anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies of the invention can be used to detect and isolate CARD-3, CARD-4, CARD-5, or CARD-6 proteins and modulate CARD-3, CARD-4, CARD-5, or CARD-6 activity.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

#### A. Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to CARD-3, CARD-4, CARD-5, or CARD-6 proteins or biologically active portions thereof or have a stimulatory or inhibitory effect on, for example, CARD-3, CARD-4, CARD-5, or CARD-6 expression or CARD-3, CARD-4, CARD-5, or CARD-6 activity. An example of a biologically active portion of human CARD-4 is amino acids 1-145 encoding the CARD domain which is sufficient to exhibit CARD-3-binding activity as described in Example 7. Amino acids 406-953 of human CARD-4L comprising the leucine rich repeat domain represent a biologically active portion of CARD-4L because they possess

hNUDC-binding activity as described in Example 8. An example of a biologically active portion of human CARD-5 is amino acids 111-881 (SEQ ID NO:58) encoding the CARD domain.

Among the screening assays provided by the invention are screening  
5 to identify molecules that prevent the dimerization of a CARD-containing polypeptide of the invention, screening to identify molecules which block the binding of a CARD containing polypeptide to a CARD-containing polypeptide of the invention (e.g., CARD-4), screening to identify a competitive inhibitor of the binding of a nucleotide to the nucleotide binding site of a CARD-containing  
10 polypeptide of the invention, e.g., human CARD-4L, screening to identify compounds which block the interaction between the leucine-rich repeat of a CARD-containing polypeptide of the invention and a ligand which binds to the leucine-rich repeat.

For CARD-6 screening assays can be used to identify molecules  
15 which modulate a CARD-6 mediated increase in transcription of genes having an AP-1 OR  $\text{nf-}\kappa\text{B}$  binding site. For example, expression of a reporter under the control of NF- $\kappa\text{B}$  (or AP-1) is measured in the presence and absence of a candidate molecule and in the presence and absence of CARD-6 to identify those molecules which alter expression of the reporter in a CARD-6 dependent manner.  
20 In addition, screening assays can be used to identify molecules which modulate a CARD-6 mediated increase in CHOP phosphorylation. For example, the expression of a reporter gene under the control of CHOP is measured in the presence and absence of a candidate small molecule and in the presence and absence of CARD-6 to identify those molecules which alter expression of the  
25 reporter in a CARD-6 dependent manner. A screening assay can be carried out to identify molecules which modulate the CARD-6 mediated increase in CHOP phosphorylation. For example, CHOP phosphorylation is measured in the presence and absence of a candidate molecule and in the presence and absence of CARD-6. Phosphorylation of CHOP can be measured using an antibody which  
30 binds to phosphorylated CHOP, but not to non-phosphorylated CHOP.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a CARD-3, CARD-4, CARD-5, or CARD-6 proteins or polypeptides or biologically active portions thereof. The test compounds of the present invention can be obtained  
5 using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to  
10 peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145). Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci.*  
15 *USA* 91:11422; Zuckermann et al. (1994). *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten  
20 (1992) *Bio/Techniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Patent No. 5,223,409), spores (Patent Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla et al.  
25 (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382; and Felici (1991) *J. Mol. Biol.* 222:301-310).

Determining the ability of the test compound to modulate the activity of CARD-3, CARD-4, CARD-5, or CARD-6 or a biologically active portion thereof can be accomplished, for example, by determining the ability of the  
30 CARD-3, CARD-4, CARD-5, or CARD-6 protein to bind to or interact with a

CARD-3, CARD-4, CARD-5, or CARD-6 target molecule. As used herein, a "target molecule" is a molecule with which a CARD-3, CARD-4, CARD-5, or CARD-6 protein binds or interacts in nature, for example, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A

5 CARD-3, CARD-4, CARD-5, or CARD-6 target molecule can be a non-CARD-3, CARD-4, CARD-5, or CARD-6 molecule or a CARD-3, CARD-4, CARD-5, or CARD-6 protein or polypeptide of the present invention. In one embodiment, a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule is a component of an apoptotic signal transduction pathway, e.g., CARD-3 and CARD-4. The target,

10 for example, can be a second intracellular protein which has catalytic activity or a protein which facilitates the association of downstream signaling molecules with CARD-3, CARD-4, CARD-5, or CARD-6. In another embodiment, CARD-3, CARD-4, CARD-5, or CARD-6 target molecules include CARD-3 because CARD-3 was found to bind to CARD-4 (Examples 7 and 12) and hNUDC

15 because hNUDC was found to bind to CARD-4 (Example 8).

Determining the ability of the test compound to modulate the activity of CARD-3, CARD-4, CARD-5, or CARD-6 or a biologically active portion thereof can be accomplished, for example, by determining the ability of the CARD-3, CARD-4, CARD-5, or CARD-6 protein to bind to or interact with any

20 of the specific proteins listed in the previous paragraph as CARD-3, CARD-4, CARD-5, or CARD-6 target molecules. In another embodiment, CARD-3, CARD-4, CARD-5, or CARD-6 target molecules include all proteins that bind to a CARD-3, CARD-4, CARD-5, or CARD-6 protein or a fragment thereof in a two-hybrid system binding assay which can be used without undue

25 experimentation to isolate such proteins from cDNA or genomic two-hybrid system libraries. For example, Example 7 describes the use of the CARD-4 CARD domain region to identify CARD-3 in a two-hybrid screen and Example 8 describes the use of the CARD-4 leucine rich repeat domain region to identify hNUDC in a two-hybrid screen. The binding assays described in this section can

30 be cell-based or cell free (described subsequently).

Determining the ability of the CARD-3, CARD-4, CARD-5, or CARD-6 protein to bind to or interact with a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule can be accomplished by one of the methods described above for determining direct binding. In an embodiment, determining the ability of the CARD-3, CARD-4, CARD-5, or CARD-6 protein to bind to or interact with a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (e.g., intracellular  $\text{Ca}^{2+}$ , diacylglycerol, IP3, etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (e.g., a CARD-3, CARD-4, CARD-5, or CARD-6-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g. luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

For example, in Example 12 CARD-4 is shown to bind to CARD-3 and in Example 10, by monitoring a cellular response, CARD-4 is shown to enhance caspase 9 activity, cell death or apoptosis. Because CARD-3 and CARD-4 enhance caspase 9 activity, activity can be monitored by assaying the caspase 9-mediated apoptosis cellular response or caspase 9 enzymatic activity. In addition, and in another embodiment, genes induced by CARD-3, CARD-4, CARD-5, or CARD-6 expression can be identified by expressing CARD-3, CARD-4, CARD-5, or CARD-6 in a cell line and conducting a transcriptional profiling experiment wherein the mRNA expression patterns of the cell line transformed with an empty expression vector and the cell line transformed with a CARD-3, CARD-4, CARD-5, or CARD-6 expression vector are compared. The promoters of genes induced by CARD-3, CARD-4, CARD-5, or CARD-6 expression can be operatively linked to reporter genes suitable for screening such as luciferase, secreted alkaline phosphatase, or beta-galactosidase and the resulting constructs could be introduced into appropriate expression vectors. A recombinant cell line containing CARD-3, CARD-4, CARD-5, or CARD-6 and transfected with an

expression vector containing a CARD-3, CARD-4, CARD-5, or CARD-6 responsive promoter operatively linked to a reporter gene can be used to identify test compounds that modulate CARD-3, CARD-4, CARD-5, or CARD-6 activity by assaying the expression of the reporter gene in response to contacting the

5 recombinant cell line with test compounds. CARD-3, CARD-4, CARD-5, or CARD-6 agonists can be identified as increasing the expression of the reporter gene and CARD-3, CARD-4, CARD-5, or CARD-6 antagonists can be identified as decreasing the expression of the reporter gene.

In another embodiment of the invention, the ability of a test

10 compound to modulate the activity of CARD-3, CARD-4, or biologically active portions thereof can be determined by assaying the ability of the test compound to modulate CARD-3, CARD-4, CARD-5, or CARD-6-dependent pathways or processes where the CARD-3, CARD-4, CARD-5, or CARD-6 target proteins that mediate the CARD-3, CARD-4, CARD-5, or CARD-6 effect are known or

15 unknown. Potential CARD-3, CARD-4, CARD-5, or CARD-6-dependent pathways or processes include, but are not limited to, the modulation of cellular signal transduction pathways and their related second messenger molecules (e.g., intracellular  $\text{Ca}^{2+}$ , diacylglycerol, IP3, cAMP etc.), cellular enzymatic activities, cellular responses (e.g., cell survival, cellular differentiation, or cell

20 proliferation), or the induction or repression of cellular or heterologous mRNAs or proteins. CARD-3, CARD-4, CARD-5, or CARD-6-dependent pathways or processes could be assayed by standard cell-based or cell free assays appropriate for the specific pathway or process under study. For example, Example 9 describes how expression of CARD-4S or CARD-4L in 293T cells induces the

25 NF- $\kappa$ B pathway as determined by the measurement of a cotransfected NF- $\kappa$ B pathway luciferase reporter gene. In another embodiment, cells cotransfected with CARD-4 and the NF- $\kappa$ B luciferase reporter gene could be contacted with a test compound and test compounds that block CARD-4 activity could be identified by their reduction of CARD-4-dependent NF- $\kappa$ B pathway luciferase

30 reporter gene expression. Test compounds that agonize CARD-4 would be



expected to increase reporter gene expression. In another embodiment, CARD-4 could be expressed in a cell line and the recombinant CARD-4-expressing cell line could be contacted with a test compound. Test compounds that inhibit CARD-4 activity could be identified by their reduction of CARD-4-dependent NF- $\kappa$ B pathway stimulation as measured by the assay of a NF- $\kappa$ B pathway reporter gene, NF- $\kappa$ B nuclear localization, I $\kappa$ B phosphorylation or proteolysis, or other standard assays for NF- $\kappa$ B pathway activation known to those skilled in the art.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof. Binding of the test compound to the CARD-3, CARD-4, CARD-5, or CARD-6 protein can be determined either directly or indirectly as described above. In one embodiment, a competitive binding assay includes contacting the CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof with a compound known to bind CARD-3, CARD-4, CARD-5, or CARD-6 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a CARD-3, CARD-4, CARD-5, or CARD-6 protein, wherein determining the ability of the test compound to interact with a CARD-3, CARD-4, CARD-5, or CARD-6 protein comprises determining the ability of the test compound to preferentially bind to CARD-3, CARD-4, CARD-5, or CARD-6 or biologically active portion thereof as compared to the known binding compound.

In another embodiment, an assay is a cell-free assay comprising contacting CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof.

Determining the ability of the test compound to modulate the activity of CARD-3, CARD-4, CARD-5, or CARD-6 can be accomplished, for example, by determining the ability of the CARD-3, CARD-4, CARD-5, or CARD-6 protein to bind to a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule by one of  
5 the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of CARD-3, CARD-4, CARD-5, or CARD-6 can be accomplished by determining the ability of the CARD-3, CARD-4, CARD-5, or CARD-6 protein to further modulate a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule.  
10 For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the CARD-3, CARD-4, CARD-5, or CARD-6 protein or biologically active portion thereof with a known compound which binds CARD-3, CARD-4,  
15 CARD-5, or CARD-6 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a CARD-3, CARD-4, CARD-5, or CARD-6 protein, wherein determining the ability of the test compound to interact with a CARD-3, CARD-4, CARD-5, or CARD-6 protein comprises determining the ability of the CARD-  
20 3, CARD-4, CARD-5, or CARD-6 protein to preferentially bind to or modulate the activity of a CARD-3, CARD-4, CARD-5, or CARD-6 target molecule. The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-associated form of CARD-3, CARD-4, CARD-5, or CARD-6. A membrane-associated form of CARD-3, CARD-4, CARD-5, or  
25 CARD-6 refers to CARD-3, CARD-4, CARD-5, or CARD-6 that interacts with a membrane-bound target molecule. In the case of cell-free assays comprising the membrane-associated form of CARD-3, CARD-4, CARD-5, or CARD-6, it may be desirable to utilize a solubilizing agent such that the membrane-associated form of CARD-3, CARD-4, CARD-5, or CARD-6 is maintained in solution.  
30 Examples of such solubilizing agents include non-ionic detergents such as n-

octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)<sub>n</sub>, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO),  
5 or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either CARD-3, CARD-4, CARD-5, or CARD-6 or its target molecule to facilitate separation of complexed  
10 from uncomplexed forms of one or both of the proteins. as well as to accommodate automation of the assay. Binding of a test compound to CARD-3, CARD-4, CARD-5, or CARD-6, or interaction of CARD-3, CARD-4, CARD-5, or CARD-6 with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the  
15 reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/CARD-3, CARD-4, CARD-5, or CARD-6 fusion proteins or glutathione-S-transferase/target fusion proteins can  
20 be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, MO) or glutathione derivatized microtitre plates. which are then combined with the test compound or the test compound and either the non-adsorbed target protein or CARD-3, CARD-4, CARD-5, or CARD-6 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological  
25 conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of CARD-3, CARD-4, CARD-5, or CARD-6 binding or  
30 activity determined using standard techniques. In an alternative embodiment,

MYC or HA epitope tag CARD-3 or CARD-4 fusion proteins or MYC or HA epitope tag target fusion proteins can be adsorbed onto anti-MYC or anti-HA antibody coated microbeads or onto anti-MYC or anti-HA antibody coated microtitre plates, which are then combined with the test compound or the test  
5 compound and either the non-adsorbed target protein or CARD-3 or CARD-4 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined  
10 either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of CARD-3 or CARD-4 binding or activity determined using standard techniques. Example 12 describes an HA epitope tagged CARD-4 protein that physically interacts in a coimmunoprecipitation assay with MYC epitope tagged CARD-3. In an  
15 embodiment of the invention, HA epitope tagged CARD-4 could be used in combination with MYC epitope CARD-3 in the sort of protein-protein interaction assay described earlier in this paragraph.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, CARD-3, CARD-4,  
20 CARD-5, or CARD-6 or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated CARD-3 or CARD-4 or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates  
25 (Pierce Chemical). Alternatively, antibodies reactive with CARD-3, CARD-4, CARD-5, CARD-6 or target molecules but which do not interfere with binding of the protein to its target molecule can be derivatized to the wells of the plate, and unbound target or protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-  
30 immobilized complexes and epitope tag immobilized complexes, include

immunodetection of complexes using antibodies reactive with the CARD-3, CARD-4, CARD-5, or CARD-6 or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the CARD-3, CARD-4, CARD-5, CARD-6 or target molecule.

5           In another embodiment, modulators of CARD-3, CARD-4, CARD-5, or CARD-6 expression are identified in a method in which a cell is contacted with a candidate compound and the expression of the CARD-3, CARD-4, CARD-5, or CARD-6 promoter, mRNA or protein in the cell is determined. The level of expression of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or  
10 protein in the presence of the candidate compound is compared to the level of expression of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of CARD-3, CARD-4, CARD-5, or CARD-6  
15 expression based on this comparison. For example, when expression of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein expression. Alternatively, when expression of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein  
20 is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein expression. The level of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or protein expression in the cells can be determined by methods described herein for detecting CARD-  
25 3, CARD-4, CARD-5, or CARD-6 mRNA or protein. The activity of the CARD-3, CARD-4, CARD-5, or CARD-6 promoter can be assayed by linking the CARD-3, CARD-4, CARD-5, or CARD-6 promoter to a reporter gene such as luciferase, secreted alkaline phosphatase, or beta-galactosidase and introducing the resulting construct into an appropriate vector, transfecting a host cell line, and  
30 measuring the activity of the reporter gene in response to test compounds. For

example, two CARD-4-specific mRNAs were detected in a Northern blotting experiment, one of 4.6 kilobases and the other of 6.5-7.0 kilobases (Example 11). In Example 11, CARD-4-specific mRNA species were found to be widely distributed in the tissues and cell lines studied.

5           In yet another aspect of the invention, the CARD-3, CARD-4, CARD-5, or CARD-6 proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-10 1696; and PCT Publication No. WO 94/10300), to identify other proteins, which bind to or interact with CARD-3, CARD-4, CARD-5, or CARD-6 ("CARD-3, CARD-4, CARD-5, or CARD-6-binding proteins" or "CARD-3, CARD-4, CARD-5, or CARD-6-bp") and modulate CARD-3, CARD-4, CARD-5, or CARD-6 activity. Such CARD-3, CARD-4, CARD-5, or CARD-6-binding  
15 proteins are also likely to be involved in the propagation of signals by the CARD-3, CARD-4, CARD-5, or CARD-6 proteins as, for example, upstream or downstream elements of the CARD-3, CARD-4, CARD-5, or CARD-6 pathway. For example, Example 7 describes the construction of a two-hybrid screening bait construct including human CARD-4L amino acids 1-145 comprising the CARD  
20 domain and the use of this bait construct to screen human mammary gland and prostate gland two-hybrid libraries resulting in the identification of human CARD-3 as a CARD-4 interacting protein. In another example, Example 8 describes the construction of a two-hybrid screening bait construct including human CARD-4 amino acids 406-953 comprising the LRR domain and the use of  
25 this bait construct to screen a human mammary gland two-hybrid libraries resulting in the identification of hNUDC as a CARD-4 interacting protein.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one  
30 construct, the gene that codes for CARD-3, CARD-4, CARD-5, or CARD-6 is

fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, in vivo, forming an CARD-3, CARD-4, CARD-5, or CARD-6-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with CARD-3, CARD-4, CARD-5, or CARD-6.

In an embodiment of the invention, the ability of a test compound to modulate the activity of CARD-3, CARD-4, CARD-5, or CARD-6, or a biologically active portion thereof can be determined by assaying the ability of the test compound to block the binding of CARD-3, CARD-4, CARD-5, or CARD-6 to its target proteins in a two-hybrid system assay. Example 7 describes a two-hybrid system assay for the interaction between CARD-3 and CARD-4 and Example 8 describes a two-hybrid system assay for the interaction between CARD-4 and its target protein hNUDC. To screen for test compounds that block the interaction between CARD-3 and CARD-4 and their target proteins, which include but are not limited to CARD-3, CARD-4, and hNUDC, a yeast two-hybrid screening strain coexpressing the interacting bait and prey constructs, for example, a CARD-4 bait construct and a CARD-3 prey construct as described in Example 7, is contacted with the test compound and the activity of the two-hybrid system reporter gene, usually HIS3, lacZ, or URA3 is assayed. If the strain remains viable but exhibits a significant decrease in reporter gene activity, this would indicate that the test compound has inhibited the interaction between the bait and prey proteins. This assay could be automated for high throughput

drug screening purposes. In another embodiment of the invention, CARD-3, CARD-4, CARD-5, or CARD-6 and their target proteins could be configured in the reverse two-hybrid system (Vidal et al. (1996) Proc. Natl. Acad. Sci. USA 93:10321-6 and Vidal et al. (1996) Proc. Natl. Acad. Sci. USA 93:10315-20) designed specifically for efficient drug screening. In the reverse two-hybrid system, inhibition of a CARD-3 or CARD-4 physical interaction with a target protein would result in induction of a reporter gene in contrast to the normal two-hybrid system where inhibition of CARD-3, CARD-4, CARD-5, or CARD-6 physical interaction with a target protein would lead to reporter gene repression.

10 The reverse two-hybrid system is preferred for drug screening because reporter gene induction is more easily assayed than report gene repression.

Alternative embodiments of the invention are proteins found to physically interact with proteins that bind to CARD-3, CARD-4, CARD-5, or CARD-6. CARD-3, CARD-4, CARD-5, or CARD-6 interactors, including but not limited to hNUDC and CARD-3, could be configured into two-hybrid system baits and used in two-hybrid screens to identify additional members of the CARD-3, CARD-4, CARD-5, or CARD-6 pathway. The interactors of CARD-3, CARD-4, CARD-5, or CARD-6 interactors identified in this way could be useful targets for therapeutic intervention in CARD-3, CARD-4, CARD-5, or CARD-6 related diseases and pathologies and an assay of their enzymatic or binding activity could be useful for the identification of test compounds that modulate CARD-3, CARD-4, CARD-5, or CARD-6 activity.

15  
20  
25

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

#### B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated

30



with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

These applications are described in the subsections below.

1. Chromosome Mapping

- 5           Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. Accordingly, CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules described herein or fragments thereof, can be used to map the location of CARD-3, CARD-4, CARD-5, or CARD-6 genes on a chromosome.
- 10   The mapping of the CARD-3, CARD-4, CARD-5, or CARD-6 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

- Briefly, CARD-3, CARD-4, CARD-5, or CARD-6 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the CARD-3, CARD-4, CARD-5, or CARD-6 sequences.
- 15   Computer analysis of CARD-3, CARD-4, CARD-5, or CARD-6 sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene
- 20   corresponding to the CARD-3, CARD-4, CARD-5, or CARD-6 sequences will yield an amplified fragment. For example, in Example 6, human CARD-4-specific PCR primers were used to screen DNAs from a somatic cell hybrid panel showing that human CARD-4 maps to chromosome 7 close to the SHGC-31928
- 25   genetic marker.

- Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells
- 30   cannot grow, because they lack a particular enzyme, but human cells can, the one

human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing  
5 easy mapping of individual genes to specific human chromosomes. (D'Eustachio et al. (1983) Science 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for  
10 assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the CARD-3, CARD-4, CARD-5, or CARD-6 sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a  
15 CARD-3, CARD-4, CARD-5, or CARD-6 sequence to its chromosome include in situ hybridization (described in Fan et al. (1990) Proc. Natl. Acad. Sci. USA 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence in situ hybridization (FISH) of a DNA sequence to a  
20 metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands  
25 develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000  
30 bases will suffice to get good results at a reasonable amount of time. For a

review of this technique, see Verma et al., (Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988)).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

10 Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, e.g., Egeland et al. (1987) Nature. 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the CARD-3, CARD-4, CARD-5, or CARD-6 gene can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

## 2. Tissue Typing

The CARD-3, CARD-4, CARD-5, or CARD-6 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of  
5 restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification  
10 difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the CARD-3,  
15 CARD-4, CARD-5, or CARD-6 sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will  
20 have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The CARD-3, CARD-4, CARD-5, or CARD-6 sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these  
25 sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of  
30 polymorphisms occur in the noncoding regions, fewer sequences are necessary to

differentiate individuals. The noncoding sequences of SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:25, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:51, SEQ ID NO:54, and SEQ ID NO:60 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000  
5 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3, SEQ ID NO:9, SEQ ID NO:27, SEQ ID NO:50, SEQ ID NO:53, SEQ ID NO:56, and SEQ ID NO:62 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

10 If a panel of reagents from CARD-3, CARD-4, CARD-5, or CARD-6 sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue  
15 samples.

### 3. Use of Partial Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying,  
20 for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the  
25 biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA  
30 sequence that is unique to a particular individual). As mentioned above, actual

base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments.

Sequences targeted to noncoding regions of SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:25, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:51, SEQ ID NO: 54, and SEQ ID NO:60 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the CARD-3, CARD-4, CARD-5, or CARD-6 sequences or portions thereof, e.g., fragments derived from the noncoding regions of SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:25, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48, SEQ ID NO:51, SEQ ID NO:54, and SEQ ID NO:60 which have a length of at least 20 or 30 bases.

The sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an in situ hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such CARD-3, CARD-4, CARD-5, or CARD-6 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, e.g., CARD-3, CARD-4, CARD-5, or CARD-6 primers or probes can be used to screen tissue culture for contamination (i.e., screen for the presence of a mixture of different types of cells in a culture).

### C. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining CARD-3, CARD-4,

CARD-5, or CARD-6 protein and/or nucleic acid expression as well as CARD-3, CARD-4, CARD-5, or CARD-6 activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with CARD-3, CARD-4, CARD-5, or CARD-6 protein, nucleic acid expression or activity. For example, mutations in a CARD-3, CARD-4, CARD-5, or CARD-6 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with CARD-3, CARD-4, CARD-5, or CARD-6 protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining CARD-3, CARD-4, CARD-5, or CARD-6 protein, nucleic acid expression or CARD-3, CARD-4, CARD-5, or CARD-6 activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs or other compounds) on the expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 in clinical trials.

These and other agents are described in further detail in the following sections.

#### 1. Diagnostic Assays

An exemplary method for detecting the presence or absence of CARD-3, CARD-4, CARD-5, or CARD-6 in a biological sample involves

obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes CARD-3, CARD-4, CARD-5, or CARD-6 protein such that the presence  
5 of CARD-3, CARD-4, CARD-5, or CARD-6 is detected in the biological sample.

An agent for detecting CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to CARD-3, CARD-4, CARD-5, or CARD-6 mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length CARD-3, CARD-4, CARD-5, or CARD-  
10 6 nucleic acid, such as the nucleic acid of SEQ ID NO:1 or 3, SEQ ID NO:7 or 9, SEQ ID NO:25 or 27, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:48 or 50, SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:54 or SEQ ID NO:56, SEQ ID NO:60 or 62, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to  
15 specifically hybridize under stringent conditions to mRNA or genomic DNA, or a human CARD-4 splice variant such as the nucleic acid of SEQ ID NO:38 or SEQ ID NO:40. Other suitable probes for use in the diagnostic assays of the invention are described herein. For example, Example 11 describes the use of a nucleic acid probe to detect CARD-4 mRNAs in human tissues and cell lines and the  
20 probe used in this experiment could be used for a diagnostic assay.

An agent for detecting CARD-3, CARD-4, CARD-5, or CARD-6 protein can be an antibody capable of binding to CARD-3, CARD-4, CARD-5, or CARD-6 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. For example, polypeptides  
25 corresponding to amino acids 128-139 and 287-298 of human CARD-4L were used to immunize rabbits and produce polyclonal antibodies that specifically recognize human CARD-4L. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')<sub>2</sub>) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling  
30 (i.e., physically linking) a detectable substance to the probe or antibody, as well



as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with

5 fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect CARD-3, CARD-4, CARD-5, or CARD-6 mRNA, protein, or genomic DNA in a biological sample in vitro as well as in

10 vivo. For example, in vitro techniques for detection of CARD-3, CARD-4, CARD-5, or CARD-6 mRNA include Northern hybridizations and in situ hybridizations. For example, Example 11 contains the use of a human CARD-4L nucleic acid probe for a Northern blotting analysis of mRNA species encoded by human CARD-4L detected in RNA samples from human tissues and cell lines.

15 In vitro techniques for detection of CARD-3 or CARD-4 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. In vitro techniques for detection of CARD-3, CARD-4, CARD-5, or CARD-6 genomic DNA include Southern hybridizations. Furthermore, in vivo techniques for detection of CARD-3, CARD-4, CARD-5, or

20 CARD-6 protein include introducing into a subject a labeled anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules

25 from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a

30 control biological sample from a control subject. contacting the control sample

with a compound or agent capable of detecting CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA, or genomic DNA, such that the presence of CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of CARD-3, CARD-4,  
5 CARD-5, or CARD-6 protein, mRNA or genomic DNA in the control sample with the presence of CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of CARD-3, CARD-4, CARD-5, or CARD-6 in a biological sample (a test sample).  
10 Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of CARD-3, CARD-4, CARD-5, or CARD-6 (e.g., an immunological disorder). For example, the kit can comprise a labeled compound or agent capable of detecting CARD-3, CARD-4, CARD-5, or CARD-6 protein or mRNA in a biological sample and  
15 means for determining the amount of CARD-3, CARD-4, CARD-5, or CARD-6 in the sample (e.g., an anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibody or an oligonucleotide probe which binds to DNA encoding CARD-3, CARD-4, CARD-5, or CARD-6, e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:38, SEQ ID NO:40, SEQ  
20 ID NO:42, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:60, or SEQ ID NO:62). Kits may also include instruction for observing that the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of CARD-3, CARD-4, CARD-5, or CARD-6 if the amount of CARD-3, CARD-4, CARD-5,  
25 or CARD-6 protein or mRNA is above or below a normal level.

For antibody-based kits, the kit may comprise, for example: (1) a first antibody (e.g., attached to a solid support) which binds to CARD-3, CARD-4, CARD-5, or CARD-6 protein; and, optionally, (2) a second, different antibody which binds to CARD-3, CARD-4, CARD-5, or CARD-6 protein or the first  
30 antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit may comprise, for example: (1) a oligonucleotide, e.g., a detectably labelled oligonucleotide, which hybridizes to a CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid sequence or (2) a pair of primers useful for amplifying a CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecule.

The kit may also comprise, e.g., a buffering agent, a preservative, or a protein stabilizing agent. The kit may also comprise components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit may also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of CARD-3, CARD-4, CARD-5, or CARD-6.

## 2. Prognostic Assays

The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with CARD-3, CARD-4, CARD-5, or CARD-6 protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. As used

herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue. Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an  
5 agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (e.g., agents of a type  
10 which decrease CARD-3, CARD-4, CARD-5, or CARD-6 activity). Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity in which a test sample is obtained and CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid is  
15 detected (e.g., wherein the presence of CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity).

The methods of the invention can also be used to detect genetic  
20 lesions or mutations in a CARD-3, CARD-4, CARD-5, or CARD-6 gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one  
25 of an alteration affecting the integrity of a gene encoding a CARD-3, CARD-4, CARD-5, or CARD-6-protein, or the mis-expression of the CARD-3, CARD-4, CARD-5, or CARD-6 gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a CARD-3, CARD-4, CARD-5, or CARD-6 gene; 2) an  
30 addition of one or more nucleotides to a CARD-3, CARD-4, CARD-5, or CARD-

6 gene; 3) a substitution of one or more nucleotides of a CARD-3, CARD-4, CARD-5, or CARD-6 gene; 4) a chromosomal rearrangement of a CARD-3, CARD-4, CARD-5, or CARD-6 gene; 5) an alteration in the level of a messenger RNA transcript of a CARD-3, CARD-4, CARD-5, or CARD-6 gene; 6) aberrant  
5 modification of a CARD-3, CARD-4, CARD-5, or CARD-6 gene, such as of the methylation pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a CARD-3, CARD-4, CARD-5, or CARD-6 gene (e.g. caused by a mutation in a splice donor or splice acceptor site); 8) a non-wild type level of a CARD-3, CARD-4, CARD-5, or CARD-6-  
10 protein; 9) allelic loss of a CARD-3, CARD-4, CARD-5, or CARD-6 gene; and 10) inappropriate post-translational modification of a CARD-3, CARD-4, CARD-5, or CARD-6-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a CARD-3, CARD-4, CARD-5, or CARD-6 gene. A biological sample is a  
15 peripheral blood leukocyte sample isolated by conventional means from a subject.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively,  
20 in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) Science 241:1077-1080; and Nakazawa et al. (1994) Proc. Natl. Acad. Sci. USA 91:360-364), the latter of which can be particularly useful for detecting point mutations in the CARD-3 or CARD-4-gene (see, e.g., Abravaya et al. (1995) Nucleic Acids Res. 23:675-682). This method can include the steps of collecting a sample of  
25 cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a CARD-3, CARD-4, CARD-5, or CARD-6 gene under conditions such that hybridization and amplification of the CARD-3, CARD-4, CARD-5, or CARD-6-gene (if present) occurs, and detecting  
30 the presence or absence of an amplification product, or detecting the size of the

amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

- 5           Alternative amplification methods include: self sustained sequence replication (Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional amplification system (Kwoh, et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi et al. (1988) Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the
- 10 detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

- In an alternative embodiment, mutations in a CARD-3, CARD-4, CARD-5, or CARD-6 gene from a sample cell can be identified by alterations in
- 15 restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence
- 20 specific ribozymes (see, e.g., U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

- In other embodiments, genetic mutations in CARD-3, CARD-4, CARD-5, or CARD-6 can be identified by hybridizing a sample and control
- 25 nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin et al. (1996) Human Mutation 7:244-255; Kozal et al. (1996) Nature Medicine 2:753-759). For example, genetic mutations in CARD-3 or CARD-4 can be identified in two-dimensional arrays containing light-generated DNA probes as described in Cronin et al. supra.
- 30 Briefly, a first hybridization array of probes can be used to scan through long

stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using  
5 smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions  
10 known in the art can be used to directly sequence the CARD-3, CARD-4, CARD-5, or CARD-6 gene and detect mutations by comparing the sequence of the sample CARD-3, CARD-4, CARD-5, or CARD-6 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) Proc. Natl. Acad. Sci. USA 74:560) or Sanger ((1977) Proc. Natl. Acad. Sci. USA 74:5463). It is also  
15 contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) Bio/Techniques 19:448), including sequencing by mass spectrometry (see, e.g., PCT Publication No. WO 94/16101; Cohen et al. (1996) Adv. Chromatogr. 36:127-162; and Griffin et al.  
20 (1993) Appl. Biochem. Biotechnol. 38:147-159).

Other methods for detecting mutations in the CARD-3 or CARD-4 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) Science 230:1242). In general, the art technique of "mismatch cleavage"  
25 starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type CARD-3, CARD-4, CARD-5, or CARD-6 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair  
30 mismatches between the control and sample strands. For instance, RNA/DNA

duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, e.g., Cotton et al (1988) Proc. Natl Acad Sci USA 85:4397; Saleeba et al (1992) Methods Enzymol. 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

10 In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in CARD-3, CARD-4, CARD-5, or CARD-6 cDNAs obtained from samples of cells. For example, the mutY enzyme  
15 of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) Carcinogenesis 15:1657-1662). According to an exemplary embodiment, a probe based on a CARD-3, CARD-4, CARD-5, or CARD-6 sequence, e.g., a wild-type CARD-3, CARD-4, CARD-5, or CARD-6 sequence, is hybridized to a cDNA or other  
20 DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, e.g., U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in CARD-3, CARD-4, CARD-5, or CARD-6 genes.  
25 For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) Proc Natl. Acad. Sci USA: 86:2766, see also Cotton (1993) Mutat. Res. 285:125-144; and Hayashi (1992) Genet Anal Tech Appl 9:73-79). Single-stranded DNA fragments of sample and control CARD-3  
30 or CARD-4 nucleic acids will be denatured and allowed to renature. The



secondary structure of single-stranded nucleic acids varies according to sequence. the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than  
5 DNA), in which the secondary structure is more sensitive to a change in sequence. In an embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) Trends Genet 7:5).

In yet another embodiment, the movement of mutant or wild-type  
10 fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al. (1985) Nature 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a  
15 further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) Biophys Chem 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization. selective  
20 amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) Nature 324:163); Saiki et al. (1989) Proc. Natl Acad. Sci USA 86:6230). Such allele specific oligonucleotides are  
25 hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant  
30 invention. Oligonucleotides used as primers for specific amplification may carry

the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al. (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition, it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al. (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a CARD-3, CARD-4, CARD-5, or CARD-6 gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which CARD-3 or CARD-4 is expressed may be utilized in the prognostic assays described herein.

### 3. Pharmacogenomics

Agents, or modulators which have a stimulatory or inhibitory effect on CARD-3, CARD-4, CARD-5, or CARD-6 activity (e.g., CARD-3, CARD-4, CARD-5, or CARD-6 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (e.g., an immunological disorder) associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 activity. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of

therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a  
5 consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of CARD-3, CARD-4, CARD-5, or CARD-6 protein, expression of CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid, or mutation content of CARD-3, CARD-4, CARD-5, or CARD-6 genes in an  
10 individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Linder (1997) Clin. Chem. 43(2):254-266.  
15 In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms.  
20 For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing  
25 enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity  
30 after taking the standard and safe dose of a drug. These polymorphisms are

expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of CARD-3, CARD-4, CARD-5, or CARD-6 protein, expression of CARD-3 or CARD-4 nucleic acid, or mutation content of CARD-3, CARD-4, CARD-5, or CARD-6 genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a CARD-3, CARD-4, CARD-5, or CARD-6 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

#### 4. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 (e.g., the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein

to increase CARD-3, CARD-4, CARD-5, or CARD-6 gene expression, protein levels, or upregulate CARD-3, CARD-4, CARD-5, or CARD-6 activity, can be monitored in clinical trials of subjects exhibiting decreased CARD-3, CARD-4, CARD-5, or CARD-6 gene expression, protein levels, or downregulated CARD-3, CARD-4, CARD-5, or CARD-6 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease CARD-3, CARD-4, CARD-5, or CARD-6 gene expression, protein levels, or downregulated CARD-3, CARD-4, CARD-5, or CARD-6 activity, can be monitored in clinical trials of subjects exhibiting increased CARD-3, CARD-4, CARD-5, or CARD-6 gene expression, protein levels, or upregulated CARD-3, CARD-4, CARD-5, or CARD-6 activity. In such clinical trials, the expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 and, preferably, other genes that have been implicated in, for example, a cellular proliferation disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, and not by way of limitation, genes, including CARD-3, CARD-4, CARD-5, or CARD-6, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) which modulates CARD-3, CARD-4, CARD-5, or CARD-6 activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of CARD-3, CARD-4, CARD-5, or CARD-6 and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of CARD-3, CARD-4, CARD-5, or CARD-6 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In an embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA, or genomic DNA in the pre-administration sample with the CARD-3, CARD-4, CARD-5, or CARD-6 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of CARD-3, CARD-4, CARD-5, or CARD-6 to lower levels than detected, i.e., to decrease the effectiveness of the agent.

#### 5. Transcriptional Profiling

The CARD-3, CARD-4, CARD-5, and CARD-6 nucleic acid molecules described herein, including small oligonucleotides, can be used in transcriptionally profiling. For example, these nucleic acids can be used to examine the expression of CARD-3, CARD-4, CARD-5, and CARD-6 in normal tissue or cells and in tissue or cells subject to a disease state, e.g., tissue or cells derived from a patient having a disease of interest or cultured cells which model or reflect a disease state of interest, e.g., cells of a cultured tumor cell line. By measuring expression of CARD-3, CARD-4, CARD-5, and CARD-6, together or

individually, a profile of expression in normal and disease states can be developed. This profile can be used diagnostically and to examine the effectiveness of a therapeutic regime.

5 C. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity, examples of which are provided herein.

10 1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity, by administering to the subject an agent which modulates CARD-3, CARD-4, CARD-5, or CARD-6 expression or  
15 at least one CARD-3, CARD-4, CARD-5, or CARD-6 activity. Subjects at risk for a disease which is caused or contributed to by aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of  
20 symptoms characteristic of the CARD-3, CARD-4, CARD-5, or CARD-6 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of CARD-3, CARD-4, CARD-5, or CARD-6 aberrancy, for example, a CARD-3, CARD-4, CARD-5, or CARD-6 agonist or CARD-3, CARD-4, CARD-5, or CARD-6 antagonist agent can be  
25 used for treating the subject. The appropriate agent can be determined based on screening assays described herein. Activities of CARD-3, CARD-4, CARD-5, or CARD-6 that could be modulated for prophylactic purposes include, but are not limited to: 1) CARD-3, CARD-4, CARD-5, or CARD-6 gene or protein expression, for example, see Example 11 for a description of the mRNA  
30 expression pattern of human CARD-4; 2) CARD-3, CARD-4, CARD-5, or

CARD-6 binding to a target protein, for example, see Examples 7, 8, and 12 for a description of proteins known to bind to CARD-3 or CARD-4; 3) CARD-4 regulation of NF- $\kappa$ B as described in Example 9; and 4) CARD-3 and CARD-4 enhancement of caspase 9 activity as described in Example 10.

5                    2.        Therapeutic Methods

Another aspect of the invention pertains to methods of modulating CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of CARD-3, CARD-4, CARD-5, or CARD-6 protein activity associated with the cell. An agent that modulates CARD-3, CARD-4, CARD-5, or CARD-6 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a CARD-3, CARD-4, CARD-5, or CARD-6 protein, a peptide, a CARD-3, CARD-4, CARD-5, or CARD-6 peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the biological activities of CARD-3, CARD-4, CARD-5, or CARD-6 protein. Examples of such stimulatory agents include active CARD-3, CARD-4, CARD-5, or CARD-6 protein and a nucleic acid molecule encoding CARD-3, CARD-4, CARD-5, or CARD-6 that has been introduced into the cell. In another embodiment, the agent inhibits one or more of the biological activities of CARD-3, CARD-4, CARD-5, or CARD-6 protein. Examples of such inhibitory agents include antisense CARD-3, CARD-4, CARD-5, or CARD-6 nucleic acid molecules and anti-CARD-3, CARD-4, CARD-5, or CARD-6 antibodies. These modulatory methods can be performed in vitro (e.g., by culturing the cell with the agent) or, alternatively, in vivo (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a CARD-3, CARD-4, CARD-5, or CARD-6 protein or nucleic acid molecule or a disorder related to CARD-3, CARD-4, CARD-5 or CARD-6 expression or activity. In one embodiment, the method involves administering an agent (e.g.,



an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. In another embodiment, the method involves administering a CARD-3, CARD-4, CARD-5, or CARD-6 protein or  
5 nucleic acid molecule as therapy to compensate for reduced or aberrant CARD-3, CARD-4, CARD-5, or CARD-6 expression or activity. Activities of CARD-3, CARD-4, CARD-5, or CARD-6 that could be modulated for therapeutic purposes include, but are not limited to, 1) CARD-3, CARD-4, CARD-5, or CARD-6 gene or protein expression, for example, see Example 11 for a description of the  
10 mRNA expression pattern of human CARD-4; 2) CARD-3, CARD-4, CARD-5, or CARD-6 binding to a target protein, for example, see Examples 7, 8, and 12 for a description of proteins known to bind to CARD-3 or CARD-4; 3) CARD-4 regulation of NF- $\kappa$ B as described in Example 9; and 4) CARD-4 enhancement of caspase 9 activity as described in Example 10.

15 Stimulation of CARD-3, CARD-4, CARD-5, or CARD-6 activity is desirable in situations in which CARD-3, CARD-4, CARD-5, or CARD-6 is abnormally downregulated and/or in which increased CARD-3, CARD-4, CARD-5, or CARD-6 activity is likely to have a beneficial effect. Conversely, inhibition of CARD-3, CARD-4, CARD-5, or CARD-6 activity is desirable in  
20 situations in which CARD-3, CARD-4, CARD-5, or CARD-6 is abnormally upregulated, e.g., in myocardial infarction, and/or in which decreased CARD-3, CARD-4, CARD-5, or CARD-6 activity is likely to have a beneficial effect. Since CARD-4 may play be involved in the processing of cytokines, inhibiting the activity or expression CARD-4 may be beneficial in patients that have  
25 aberrant inflammation.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

30

### EXAMPLES

Example 1: Isolation and Characterization of full-length Human CARD-3 and CARD-4L/S cDNAs.

A profile of known CARD domains was used to search databases of  
5 cDNA sequences and partial cDNA sequences using TBLASTN (Washington  
University; version 2.0, BLOSUM62 search matrix). This search led to the  
identification of CARD-3. Using CARD-3 to search databases of cDNA  
sequences and partial cDNA sequences, another potential CARD cDNA was  
found. This cDNA sequence was used to screen a human umbilical vein endothelial  
10 library (HUVE) and a clone containing the partial CARD-4S was identified. The  
human umbilical vein endothelial library was then rescreened using a probe  
designed against the partial CARD-4S sequence and a clone containing the  
CARD-4L sequence was identified.

15 Example 2: Characterization of CARD-3 AND CARD-4L/S Proteins.

In this example, the predicted amino acid sequences of human CARD-  
3 and CARD-4L/S proteins were compared to amino acid sequences of known  
proteins and various motifs were identified. For example, the CARD domains of  
CARD-3 and CARD-4 were aligned (Figure 7) with the CARD domains of ARC-  
20 CARD (SEQ ID NO:31), cIAP1-CARD (SEQ ID NO:32) and cIAP2-CARD  
(SEQ ID NO:33). In addition, the molecular weight of the human CARD-3 and  
CARD-4L/S proteins were predicted.

The human CARD-3 cDNA was isolated as described above (Figure  
1; SEQ ID NO:1) and encodes a 540 amino acid protein (Figure 2: SEQ ID  
25 NO:2). CARD-3 also includes one predicted kinase domain (amino acid 1 to  
amino acid 300 of SEQ ID NO:2; SEQ ID NO:4), which is followed by a  
predicted linker domain (amino acid 301 to amino acid 431 of SEQ ID NO:2;  
SEQ ID NO:5) and a predicted CARD domain (amino acid 432 to amino acid  
540 of SEQ ID NO:2; SEQ ID NO:6).

30 The human CARD-4L cDNA was isolated as described above (Figure

3; SEQ ID NO:7) and has a 2859 nucleotide open reading frame (nucleotides 245-3103 of SEQ ID NO:7; SEQ ID NO:9) which encodes a 953 amino acid protein (Figure 4; SEQ ID NO:8). CARD-4L protein has a predicted CARD domain (amino acids 15-114; SEQ ID NO:10). CARD-4L is also predicted to

5 have a nucleotide binding domain which extends from about amino acid 198 to about amino acid 397 of SEQ ID NO:8; SEQ ID NO:11, a predicted Walker Box "A", which extends from about amino acid 202 to about amino acid 209 of SEQ ID NO:8; SEQ ID NO:12, a predicted Walker Box "B", which extends from about amino acid 280 to about amino acid 284, of SEQ ID NO:8; SEQ ID NO:13,

10 a predicted kinase 1a (P-loop) domain, which extends from about amino acid 197 to about amino acid 212 of SEQ ID NO:8; SEQ ID NO:46, a predicted kinase 2 domain, which extends from about amino acid 273 to about amino acid 288 of SEQ ID NO:8; SEQ ID NO:47, a predicted kinase 3a subdomain, which extends from about amino acid 327 to about amino acid 338 of SEQ ID NO:8; SEQ ID

15 NO:14, ten predicted Leucine-rich repeats which extend from about amino acid 674 to about amino acid 950 of SEQ ID NO:8. The first Leucine-rich repeat is predicted to extend from about amino acid 674 to about amino acid 701 of SEQ ID NO:8; SEQ ID NO:15. The second Leucine-rich repeat is predicted to extend from about amino acid 702 to about amino acid 727 of SEQ ID NO:8; SEQ ID

20 NO:16. The third Leucine-rich repeat is predicted to extend from about amino acid 728 to about amino acid 754 of SEQ ID NO:8; SEQ ID NO:17. The fourth Leucine-rich repeat is predicted to extend from about amino acid 755 to about amino acid 782 of SEQ ID NO:8; SEQ ID NO:18. The fifth Leucine-rich repeat is predicted to extend from about amino acid 783 to about amino acid 810 of SEQ

25 ID NO:8; SEQ ID NO:19. The sixth Leucine-rich repeat is predicted to extend from about amino acid 811 to about amino acid 838 of SEQ ID NO:8; SEQ ID NO:20. The seventh Leucine-rich repeat is predicted to extend from about amino acid 839 to about amino acid 866 of SEQ ID NO:8; SEQ ID NO:21. The eighth Leucine-rich repeat is predicted to extend from about amino acid 867 to about

30 amino acid 894 of SEQ ID NO:8; SEQ ID NO:22. The ninth Leucine-rich repeat

is predicted to extend from about amino acid 895 to about amino acid 922 of SEQ ID NO:8; SEQ ID NO:23 and the tenth Leucine-rich repeat is predicted to extend from about amino acid 923 to about amino acid 950 of SEQ ID NO:8; SEQ ID NO:24.

5           The human partial CARD-4S cDNA isolated as described above (Figure 5; SEQ ID NO:25) encodes a 490 amino acid protein (Figure 6; SEQ ID NO:26). CARD-4S includes one predicted partial CARD domain (amino acids 1-74 of SEQ ID NO:26). CARD-4S is also predicted to have a P-Loop which extends from about amino acid 163 to about amino acid 170 of SEQ ID NO:26;  
10 SEQ ID NO:29, and a predicted Walker Box "B" which extends from about amino acid 241 to about amino acid 245 of SEQ ID NO:26; SEQ ID NO:30.

A plot showing the predicted structural features of CARD-4L is presented in Figure 8. This figure shows the predicted alpha regions (Garnier-Robinson and Chou-Fasman), the predicted beta regions (Garnier-Robinson and  
15 Chou-Fasman), the predicted turn regions (Garnier-Robinson and Chou-Fasman) and the predicted coil regions (Garnier-Robinson and Chou-Fasman). Also included in the figure is a hydrophilicity plot (Kyte-Doolittle), the predicted alpha and beta-amphatic regions (Eisenberg), the predicted flexible regions (Karplus-Schulz), the predicted antigenic index (Jameson-Wolf) and the predicted surface  
20 probability plot (Emini).

A plot showing the predicted structural features of CARD-4S is also presented in Figure 9. This figure shows the predicted alpha regions (Garnier-Robinson and Chou-Fasman), the predicted beta regions (Garnier-Robinson and Chou-Fasman), the predicted turn regions (Garnier-Robinson and Chou-Fasman)  
25 and the predicted coil regions (Garnier-Robinson and Chou-Fasman). Also included in the figure is a hydrophilicity plot (Kyte-Doolittle), the predicted alpha and beta-amphatic regions (Eisenberg), the predicted flexible regions (Karplus-Schulz), the predicted antigenic index (Jameson-Wolf) and the predicted surface probability plot (Emini).

30

The predicted MW of CARD-3 is approximately 61 kDa. The predicted MW of CARD-4L is approximately 108 kDa.

#### Example 3: Preparation of CARD-3 and CARD-4 Proteins

5           Recombinant CARD-3 and CARD-4 can be produced in a variety of expression systems. For example, the CARD-3 and CARD-4 peptides can be expressed as a recombinant glutathione-S-transferase (GST) fusion protein in *E. coli* and the fusion protein can be isolated and characterized. Specifically, as described above, CARD-3 or CARD-4 can be fused to GST and the fusion  
10 protein can be expressed in *E. coli* strain PEB199. Expression of the GST-CARD-3 or GST-CARD-4 fusion protein in PEB199 can be induced with IPTG. The recombinant fusion protein can be purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads.

#### 15 Example 4: Identification of splice variants of CARD-4.

The 5' untranslated sequence from CARD-4L was used to search databases of cDNA sequences and partial cDNA sequences using BLASTN (Washington University; version 2.0, BLOSUM62 search matrix) for additional CARD-4 cDNA clones. This search led to the identification of two cDNA  
20 clones, clone Z from a human lymph node library and the Y clone from a human brain cDNA library. Both clones were sequenced and found to represent probable splice variants of CARD-4 that encode truncated CARD-4 proteins, Y encoding a 249 amino acid protein and Z encoding a 164 amino acid protein.

Fig. 10 shows the nucleotide (SEQ ID NO:38) and Fig. 11 the predicted amino  
25 acid (SEQ ID NO:39) sequences of human CARD-4Y; Fig. 12 shows the nucleotide (SEQ ID NO:40) and Fig. 13 the amino acid (SEQ ID NO:41) sequences of human CARD-4Z; and Fig. 14 shows an alignment of the CARD-4L, CARD-4Y, and CARD-4Z amino acid sequences generated by the Clustal program using a PAM250 residue weight table.

30

Example 5: Identification of murine CARD-4.

The CARD-4 polypeptide sequence was used to search databases of cDNA sequences and partial cDNA sequences using the TBLASTN program (version 1.4, BLOSUM62 search matrix, and a word length of 3) for murine  
5 CARD-4 cDNA clones. This search led to the identification of a partial murine CARD-4 clone designated murine CARD-4L. The rapid identification of cDNA ends procedure (RACE) was applied to the 5' end of the murine CARD-4L clone to elucidate the 5' end of the murine CARD-4L cDNA. Fig. 15 shows the murine CARD-4L nucleotide sequence (SEQ ID NO:42), Figure 16 shows the murine  
10 CARD-4L amino acid sequence (SEQ ID NO:43), and Fig. 17 shows an alignment of the murine CARD-4L and human CARD-4L amino acid sequences generated by the Clustal program using a PAM250 residue weight table.

Example 6: Identification of the chromosomal location of human CARD-4.

15 To determine the chromosomal location of the human CARD-4 gene, the polymerase chain reaction carried out with human CARD-4-specific primers card4t, with the 5' to 3' sequence agaaggtctggtcggcaaa (SEQ ID NO:44), and card4k, with the 5' to 3' sequence aagccctgagtggaagca (SEQ ID NO:45), was used to screen DNAs from a commercially available somatic cell hybrid panel. This  
20 analysis showed that human CARD-4 maps to chromosome 7 close to the SHGC-31928 genetic marker.

Example 7: Identification of CARD-3 in a yeast two-hybrid screen for proteins that physically interact with the CARD domain of human CARD-4.

25 DNA encoding amino acids 1-145 of human CARD-4 comprising the CARD domain was cloned into a yeast two-hybrid screening vector to create a CARD-4,1-145-GAL4 DNA-binding domain fusion for two-hybrid screening. The CARD-4,1-145-GAL4 DNA-binding domain fusion was used to screen human mammary gland and human prostate two-hybrid libraries for gene  
30 products that could physically associate with CARD-4.1-145. Twelve library

plasmids expressing CARD4,1-145 interacting proteins were found to contain the CARD-domain containing protein CARD-3 thus establishing a direct or indirect physical interaction between CARD-4 and CARD-3.

In addition, DNA encoding amino acids 435-540 of CARD-3  
5 comprising the CARD domain of CARD-3 (SEQ ID NO:6) was cloned into a yeast two-hybrid GAL4 transcriptional activation domain fusion vector to create a CARD-3,435-540-GAL4 transcriptional activation domain fusion. To test whether the CARD domain of CARD-3 binds CARD-4, 1-145, the CARD-3,435-540-GAL4 transcriptional activation domain fusion expression vector and the  
10 CARD-4,1-145-GAL4 DNA-binding domain fusion vector were cotransformed into a two-hybrid screening *Saccharomyces cerevisiae* (yeast) strain. The resulting cotransformed yeast strain expressed the two reporter genes that indicate a physical interaction between the two hybrid proteins in the experiment, in this case, the CARD-3,435-540-GAL4 transcriptional activation domain fusion  
15 protein and the CARD-4,1-145-GAL4 DNA-binding domain fusion protein. This experiment established a physical interaction between the CARD domain of CARD-3 and the CARD domain of CARD-4.

Example 8: Identification of hNUDC in a yeast two-hybrid screen for proteins  
20 that physically interact with the LRR domain of human CARD-4.

DNA encoding amino acids 406-953 of human CARD-4L comprising the LRR domain was cloned into a yeast two-hybrid screening vector to create a CARD-4,406-953-GAL4 DNA-binding domain fusion for two-hybrid screening. The CARD-4,406-953-GAL4 DNA-binding domain fusion was used to screen a  
25 human mammary gland two-hybrid library for gene products that could physically associate with CARD-4,406-953. One library plasmid expressing a CARD-4,406-953 interacting protein was found to contain the hNUDC protein, the human ortholog of the rat NUDC protein that has been implicated in nuclear movement (Morris et al., Curr. Biol. 8:603 [1998], Morris et al., Exp. Cell Res.

30

238:23 [1998]), thus establishing a physical interaction between CARD-4 and hNUDC.

Example 9: Discovery of regulation by CARD-4 of NF- $\kappa$ B.

5           The first group of experiments described in this Example were carried out to determine if CARD-4 can activate the NF- $\kappa$ B pathway. CARD-4 regulation of the NF- $\kappa$ B pathway is of interest because the NF- $\kappa$ B pathway is involved in many diseases described in (New England Journal of Medicine 336:1066 [1997]) and (American Journal of Cardiology 76:18C [1995]) and other  
10 references known to those skilled in the art. Participation of CARD-4 in the NF- $\kappa$ B pathway would make CARD-4 an attractive target for drugs that modulate the NF- $\kappa$ B pathway for treatment of NF- $\kappa$ B pathway-dependent diseases, conditions, and biological processes.

          The first group of experiments showed specific CARD-4-mediated  
15 NF- $\kappa$ B pathway induction.

          The second group of experiments described in this Example were carried out to determine if CARD-3, the NIK serine/threonine protein kinase (Su et al., EMBO J. 16:1279 [1997]), or the signal transduction protein TRAF6 (Cao et al., Nature 383:443 [1996]), proteins known to participate in the induction of  
20 NF- $\kappa$ B (McCarthy et al., J. Biol. Chem. 273:16968 [1998]), are involved in transducing the CARD-4-dependent NF- $\kappa$ B pathway induction signal. It was found that CARD-3, NIK, and TRAF6 are all involved in transducing the CARD-4-mediated NF- $\kappa$ B pathway induction signal.

          In nine transfection experiments, 293T cells coexpressing an NF- $\kappa$ B  
25 reporter plasmid and either pCI, pCI-CARD-4L (expressing CARD-4L), pCI-CARD-4S (expressing CARD-4S), pCI-APAF1 (expressing Apaf-1), pCI-APAFS (expressing an Apaf-1 variant lacking WD repeats), pCI-CARD-4LnoCARD (expressing CARD-4L without a CARD domain), pCI-CARD4LnoLRR (expressing CARD-4L without a LRR), pCI-  
30 CARD4LCARDonly (expressing CARD-4L CARD domain only), or pCI-



CARD4NBSonly (expressing CARD-4L nucleotide binding sequence only) were created. 293T cells were plated in 6-well plates (35 mm wells) and transfected 2 days later (90% confluency) with 1  $\mu$ g of NF- $\kappa$ B luciferase reporter plasmid (pNF- $\kappa$ B-Luc, Stratagene), 200 ng of pCMV  $\beta$ -gal, 600 ng of pCI vector and 200 ng of indicated expression plasmids using SuperFect transfection reagent (Qiagen). For dominant-negative experiments, 2 ng of CARD4 expressing plasmid and 800 ng of dominant-negative plasmid were used. Cells were harvested 48 h after transfection and luciferase activity in 1000-fold diluted cell extracts was determined using the Luciferase Assay System (Promega). In addition,  $\beta$ -galactosidase activities were determined and used to normalize transfection efficiency.

Relative luciferase activity was determined at the end of the experiment to assess NF- $\kappa$ B pathway activation by the gene expressed by the pCI-based plasmid in each transfected cell line. The cell lines containing pCI, pCI-APAFS, pCI-APAF, pCI-CARD-4LnoCARD, and pCI-CARD4NBSonly had similar baseline levels of luciferase expression but the cell lines containing pCI-CARD-4L, pCI-CARD4LnoLRR, and pCI-CARD4LCARDonly had luciferase expression about nine fold elevated relative to baseline and the cell line containing pCI-CARD4S had luciferase expression sixteen fold elevated relative to baseline. This result demonstrates induction by CARD-4S and CARD-4L of the NF- $\kappa$ B pathway. This CARD-4 mediated NF- $\kappa$ B pathway induction is dependent on the CARD-4 CARD domain because the pCI-CARD-4noCARD construct expressing CARD-4 lacking its CARD domain did not induce the luciferase reporter gene and pCI-CARD4LCARDonly expressing the CARD-4 CARD domain did induce the luciferase reporter gene. Also, the CARD-4 LRR domains are not required for NF- $\kappa$ B pathway activation because pCI-CARD4LnoLRR expressing a CARD-4 mutant protein lacking LRR domains is able to induce the luciferase reporter gene. In addition, the CARD-4 NBS domain is not sufficient for NF- $\kappa$ B pathway activation because pCI-CARD4NBSonly expressing CARD-4 NBS domain is not able to induce the

luciferase reporter gene. In addition, the induction of the NF- $\kappa$ B pathway by CARD-4 is specific, as neither Apaf-expressing construct in this experiment induced luciferase activation.

In five transfection experiments, 293T cells coexpressing an NF- $\kappa$ B reporter plasmid (NF- $\kappa$ B-luciferase, Stratagene) and pCI-CARD-4L and either, no vector, pCI-TRAF6-DN (expressing a dominant negative version of TRAF-6), pCI-NIK-DN (expressing a dominant negative version of NIK kinase), pCI-CARD3CARDonly (expressing the CARD domain of CARD-3, which acts as a dominant negative version of CARD-3), or pCI-Bcl-XL (expressing the anti-apoptotic protein Bcl-XL) were created. TRAF6-DN, NIK-DN, and CARD3CARDonly are dominant negative alleles of the TRAF6, NIK, and CARD3 genes, respectively. After 48 hours, cells were lysed and the relative luciferase activity was determined (Promega Kit) to assess NF- $\kappa$ B pathway activation by the genes expressed by the one or two pCI-based plasmids in each transfected cell line. The cell lines containing pCI-CARD-4L only or pCI-CARD-4L and pCI-Bcl-XL had relative luciferase reporter gene expression of about 18 units. The cell lines containing pCI-CARD-4L and pCI-TRAF6-DN, pCI-CARD-4L and pCI-NIK-DN, or pCI-CARD-4L and pCI-CARD3CARDonly had relative luciferase reporter gene expression of about 4 units. Inhibition of CARD-4L-mediated NF- $\kappa$ B pathway induction by TRAF6-DN, NIK-DN, and CARD3CARDonly is specific as Bcl-XL did not inhibit CARD-4L-mediated NF- $\kappa$ B pathway induction.

These results demonstrate that dominant negative alleles of TRAF6, NIK and CARD-3 expressed, respectively, from pCI-TRAF6-DN, pCI-NIK-DN, and pCI-CARD3CARDonly block induction of the NF- $\kappa$ B reporter gene by CARD-4L expression (pCI-CARD-4L) and suggest that TRAF6, NIK, and CARD-3 act downstream of CARD-4L to transduce the CARD-4L NF- $\kappa$ B pathway induction stimulus.

In an additional experiment, coexpression of CARD-4 and the CARD domain of CARD-3 revealed that the CARD domain of CARD-3 functions as a

dominant negative mutant suggesting that CARD-3 is a downstream mediator of CARD-4 function.

Example 10: Discovery of CARD-4 enhancement of caspase-9 activity.

5 In ten transfection experiments, 293T cells coexpressing a beta galactosidase-expressing plasmid (pCMV  $\beta$ -gal from Stratagene) as a marker for viable cells and either pCI, pCI-CARD-3, pCI-APAF, pCI-CARD-4L, pCI-CARD-4S, pCI-CARD4LnoLRR, pCI-CARD4NBsOnly, pCI-CARD4LCARDonly, pCI-CARD-4LnoCARD or pCI-casp9 (expressing caspase-  
10 9) were created. Transfections included 400 ng of pCMV  $\beta$ -gal, 800 ng of expression plasmid, and Superfect transfection reagent from Qiagen and were carried out according to the manufacturer's directions. After 40-48 hours, cells were fixed and stained for beta-galactosidase expression and cell viability was determined by counting the number of beta galactosidase positive cells.  
15 Expression of pCI, pCI-CARD-3, pCI-APAF, pCI-CARD-4L, pCI-CARD-4S, pCI-CARD4LnoLRR, pCI-CARD4NBsOnly, pCI-CARD4LCARDonly, and pCI-CARD-4LnoCARD did not result in loss of cell viability. As expected, expression of pCI-casp9 in 293T cells resulted in a loss of viability of about 75% of the cells in the experiment.

20 It was next tested whether pCI, pCI-CARD-3, pCI-APAF, pCI-CARD-4L, pCI-CARD-4S, pCI-CARD4LnoLRR, pCI-CARD4NBsOnly, pCI-CARD4LCARDonly, or pCI-CARD-4LnoCARD would regulate caspase 9-mediated apoptosis. In nine transfection experiments, 293T cells coexpressing a beta galactosidase-expressing plasmid as a marker for viable cells, pCI-casp9,  
25 and either pCI, pCI-CARD-3, pCI-APAF, pCI-CARD-4L, pCI-CARD-4S, pCI-CARD4LnoLRR, pCI-CARD4NBsOnly, pCI-CARD4LCARDonly, and pCI-CARD-4LnoCARD were created. After 40-48 hours, cells were fixed and stained for beta-galactosidase expression and cell viability was determined by counting the number of beta galactosidase positive cells. Expression of pCI, pCI-  
30 CARD-4LnoCARD, and pCI-CARD4NBsOnly in the caspase 9-expressing 293T

cells had no effect on the caspase 9-induced apoptosis. However, pCI-CARD-3, pCI-CARD-4L, pCI-CARD-4S, pCI-CARD4LnoLRR, pCI-CARD4LCARDonly and, as expected, pCI-APAF enhanced the level of caspase 9-induced apoptosis to 20 or less beta galactosidase positive cells per experiment from about 100 beta  
5 galactosidase positive cells per experiment.

This experiment demonstrated that CARD-4 can enhance caspase 9-mediated apoptosis because coexpression of CARD-4L or CARD-4S with caspase-9 dramatically increases caspase-9 mediated apoptosis. Furthermore, the CARD-4 CARD domain (SEQ ID NO:10) is necessary and sufficient for CARD-  
10 4-mediated enhancement of caspase-9-potentiated apoptosis because CARD-4L lacking its CARD domain (pCI-CARD-4LnoCARD) does not enhance caspase-9-mediated apoptosis while the CARD-4 CARD domain expressed alone (pCI-CARD4LCARDonly) does induce caspase-9 mediated apoptosis. In addition, the LRR present in CARD-4 is not required for CARD-4 enhancement of caspase-9-  
15 mediated apoptosis because expression of a CARD-4 protein lacking the LRR (pCI-CARD4LnoLRR) still enhances caspase-9-mediated apoptosis. The CARD-4 NBS is not sufficient for CARD-4 enhancement of caspase-9-mediated apoptosis because expression of the CARD-4 NBS only (pCI-CARD4NBSonly) does not enhance caspase-9 mediated apoptosis. This experiment also  
20 demonstrates that CARD-3 can enhance caspase-9-mediated apoptosis.

As detailed below in Example 12, CARD-4 does not appear to interact directly with caspase-9, suggesting that potentiation of caspase-9 activity by CARD-4 is mediated by activation of downstream pathways.

25 Example 11: Identification and tissue distribution of mRNA species expressed by the human CARD-4 gene.

Northern analysis of mRNAs extracted from adult human tissues revealed a 4.6 kilobase mRNA band that was expressed in most tissues examined. Highest expression was observed heart, spleen, placenta and lung. CARD-4 was  
30 also observed to be expressed in fetal brain, lung, liver and kidney. Cancer cell

lines expressing the 4.6 kilobase CARD-4 mRNA include HeLa, K562, Molt4, SW480, A549 and melanoma. A larger 6.5 to 7.0 kilobase CARD-4 mRNA was expressed in heart, spleen, lung, fetal lung, fetal liver, and in the Molt4 and SW480 cell lines.

5

Example 12: Physical association of CARD-4 with CARD-3.

- CARD-4-specific PCR primers with the 3' primer encoding the HA epitope tag were used to amplify the CARD-4L gene epitope tagged with HA and this PCR product was cloned into the mammalian expression vector pCI. CARD-3-specific PCR primers with the 5' primer encoding the MYC epitope tag were used to amplify the CARD-3 gene epitope tagged with MYC and this PCR product was cloned into the mammalian expression vector pCI. CARD-3-specific PCR primers with the 5' primer encoding the MYC epitope tag were used to amplify the CARD-3 gene lacking the CARD domain (SEQ ID NO:6) epitope tagged with MYC and this PCR product was cloned into the mammalian expression vector pCI. Caspase 9-specific PCR primers with the 3' primer encoding the MYC epitope tag were used to amplify the caspase 9 gene epitope tagged with MYC and this PCR product was cloned into the mammalian expression vector pCI. In three transfection experiments, 293T cells coexpressing pCI-CARD-4LcHA and either pCI-CARD3nMYC, pCI-CARD3noCARDnMYC, or pCI-casp9cMYC were created. Cells from each transfected line were lysed and an immunoprecipitation procedure was carried out on each lysate with an anti-MYC epitope tag antibody to precipitate the CARD-4LcHA expressed by each cell line and any physically associated proteins. Immunoprecipitated proteins were separated by electrophoresis on denaturing polyacrylamide gels, transferred to nylon filters, and probed with an anti-HA epitope tag antibody in a Western blotting experiment to determine whether the MYC-tagged protein that was coexpressed with the CARD-4LcHA protein had coimmunoprecipitated with the CARD-4LcHA protein. In this experiment, CARD-3 was found to coimmunoprecipitate with CARD-4 while CARD-3

lacking its CARD domain and caspase-9 did not coimmunoprecipitate with CARD-4. This experiment demonstrates that CARD-4 and CARD-3 physically associate and that CARD-3 requires its CARD domain to associate with CARD-4. In addition, CARD-4 appears to not associate with caspase-9.

5

#### Example 13: CARD-4 Genomic Sequence

Figure 18 depicts the 32042 nucleotide genomic sequence of CARD-4 (SEQ ID NO:63). This sequence is based the CARD-4 cDNA sequence described above and a BAC sequence (DBEST Accession No. AC006027). The CARD-4 cDNA sequence described above was used to correct three errors in the BAC sequence, including one error resulting in a frameshift. The CARD-4 genomic sequence of Figure 18 includes the following introns and exons: exon 1: nucleotides 364-685, encoding amino acids 1-67 (start codon at nucleotides 485-487); intron 1: nucleotides 686-2094; exon 2: nucleotides 2095-2269, encoding amino acids 67-126; intron 2: nucleotides 2270-4365; exon 3: nucleotides 366-6190, encoding amino acids 126-734; intron 3: nucleotides 6191-9024; exon 4: nucleotides 9025-9108, encoding amino acids 734-762; intron 4: nucleotides 9109-10355; exon 5: nucleotides 10356-10439, encoding amino acids 762-790; intron 5: nucleotides 10440-11181; exon 6: nucleotides 11182-11265, encoding amino acids 790-818; intron 6: nucleotides 11266-19749; exon 7: nucleotides 19750-19833, encoding amino acids 818-846; intron 7: nucleotides 19834-21324; exon 8: nucleotides 21325-21408, encoding amino acids 846-874; intron 8: nucleotides 21409-24226; exon 9: nucleotides 24227-24310, amino acids 874-903; intron 9: nucleotides 24311-27948; exon 10: nucleotides 27949-28032, amino acids 903-930; intron 10: nucleotides 28033-31695; exon 11: nucleotides 31696-32024, encoding amino acids 930-953 (stop codon at nucleotides 31766-31768).

The introns in the CARD-4 genomic sequence contain consensus splice donor and acceptor sites (Molecular Cell Biology, Darnell et al., eds., 1996). The CARD-4 genomic sequence is useful for genetic identification

and mapping and identifying mutations, e.g., mutations is splice donor or splice acceptor sites.

Example 14: Isolation and characterization of full-length murine CARD-5 and  
5 human CARD-5

The amino acid sequence of the CARD domain of RAID (amino acids 1 to 94) was used to search a proprietary murine cDNA sequence database using the BLASTX program with the BLOSUM62 matrix and a protein word length of three. This search led to the identification of a murine clone,  
10 jtmma010ht2, present in a coronary artery smooth muscle cell library. This clone encodes a protein designated CARD-5. The 761 nucleotide murine CARD-5 cDNA of SEQ ID NO:60 has a 579 nucleotide open reading frame (SEQ ID NO:62) encoding a 193 amino acid protein (SEQ ID NO:61). The cDNA and protein sequences of murine CARD-5 are shown in Figure 19.

15 Murine CARD-5 is predicted to be an intracellular protein having a molecular weight of 21.4 kDa prior to post-translational modification.

Figure 20 depicts a hydropathy plot of murine CARD-5. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues  
20 (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

The murine CARD-5 nucleotide sequence was used to search a proprietary database of human cDNA sequences. This search led to the identification of a human CARD-5 cDNA clone, jthza027g11t1, present in a  
25 testes library.

The 740 nucleotide murine CARD-5 cDNA of SEQ ID NO:48 has a 585 nucleotide open reading frame (SEQ ID NO:50) encoding a 195 amino acid protein (SEQ ID NO:49). The cDNA and protein sequences of human CARD-5 are shown in Figure 21.

30

Human CARD-5 is predicted to be an intracellular protein having a molecular weight of 21.6 kDa prior to post-translational modification.

Figure 22 depicts a hydropathy plot of human CARD-5. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 23 depicts an alignment of the cDNA sequences of murine (SEQ ID NO:60) and human (SEQ ID NO:48) CARD-5. In this alignment the sequences are 68.2% identical. Figure 24 depicts an alignment of the amino acid sequences of murine (SEQ ID NO:61) and human (SEQ ID NO:49) CARD-5. In this alignment the sequences are 71.8% identical.

Both murine and human CARD-5 include a CARD domain. The CARD domain of murine CARD-5 extends from amino acid 110 to 179 of SEQ ID NO:61 (SEQ ID NO:57). The CARD domain of human CARD-5 extends from amino acid 111 to 181 of SEQ ID NO:49 (SEQ ID NO:58). Figure 27 depicts an alignment of the CARD domains of murine CARD-5 (SEQ ID NO:57), human CARD-5 (SEQ ID NO:58), and RAIDD (SEQ ID NO:70).

Example 15: Isolation and Characterization of full-length rat CARD-6 and human CARD-6

A generalized CARD domain model was used to search a proprietary rat cDNA sequence database. This search led to the identification of a rat cDNA clone present in a sciatic nerve cDNA library. This clone encodes a protein designated CARD-6. The 5252 nucleotide rat CARD-6 cDNA of SEQ ID NO:51 has a 2715 nucleotide open reading frame (SEQ ID NO:53) encoding a 905 amino acid protein (SEQ ID NO:52). The cDNA and protein sequences of rat CARD-6 are shown in Figure 25.

Rat CARD-6 is predicted to be an intracellular protein having a molecular weight of 100.2 kDa prior to post-translational modification.



Figure 26 depicts a hydropathy plot of rat CARD-6. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Rat CARD-6 contains a CARD domain which extends from amino acid 1 to amino acid 108 of SEQ ID NO:52 (SEQ ID NO:59). Rat CARD-6 also has a proline-rich c-terminus which extends from amino acid 698 to amino acid 905 of SEQ ID NO:52 (SEQ ID NO:65). This proline-rich domain includes five putative SH3 binding sites. These binding sites have the sequence PXXP and are located at amino acids 710 to 713 (PAHP), 806 to 809 (PLRP), 819 to 822 (PIPP), 857 to 860 (PPHP), and 881 to 884 (PSQP) of SEQ ID NO:52.

The rat CARD-6 cDNA sequence described above was used to search a proprietary sequence database. This search led to the identification of a clone from a human muscle cell library encoding a carboxy-terminal portion of human CARD-6. A probe designed based on the sequence of this clone was used to screen a human adrenal gland library. This screening led to the identification of a clone encoding an amino-terminal portion of human CARD-6. The clone encoding an amino terminal portion of human CARD-6 contains a region encoding a CARD domain. This CARD domain-encoding sequence was used to screen a proprietary database. This screening led to the identification of a clone, jthAb086d02, present in an adrenal gland library, which encodes full length human CARD-6.

The 4244 nucleotide human CARD-6 cDNA of SEQ ID NO:54 has a 3111 nucleotide open reading frame (SEQ ID NO:56) encoding a 1037 amino acid protein (SEQ ID NO:55). The cDNA and protein sequences of human CARD-6 are shown in Figure 28.

N-glycosylation sites are present at amino acids 49-52, 415-418, and 812-815 of SEQ ID NO:55. Human CARD-6 contains cAMP and cGMP-dependent protein kinase phosphorylation sites at amino acids 151-154 and 429-

432 of SEQ ID NO:55. Protein kinase C phosphorylation sites are present at amino acids 34-36, 57-59, 135-137, 154-156, 161-163, 298-300, 339-341, 346-348, 443-445, 664-666, 693-695, 746-748, 882-884, 905-907 and 951-953 of SEQ ID NO:55. Casein kinase II phosphorylation sites are present at amino acids  
5 6-9, 28-31, 40-43, 112-115, 135-138, 154-157, 278-281, 321-324, 339-342, 354-357, 642-645, 670-673, and 707-710 of SEQ ID NO:55. Tyrosine kinase phosphorylation sites are present at amino acids 37-34 and 163-169 of SEQ ID NO:55. An ATP/GTP-binding site motif A (P-loop) site is present at amino acids 775-782 of SEQ ID NO:55.

10           Figure 29 depicts a hydropathy plot of human CARD-6. Relatively hydrophobic regions are above the horizontal line, and relatively hydrophilic regions are below the horizontal line. Cysteine residues are indicated by short vertical lines just below the hydropathy trace.

          Human CARD-6 is predicted to have a molecular weight of 116.5 kD  
15 before post-translational modification.

          Human CARD-6 includes a CARD domain at amino acids 5-92 of SEQ ID NO:55 (SEQ ID NO:64). Figure 30 depicts an alignment of the CARD domain domain of human CARD-6 and a consensus CARD domain derived from a hidden Markov model (SEQ ID NO:67).

20           Northern blot analysis of rat CARD-6 expression revealed that CARD-6 is expressed at a high level in the heart (6.5 kb transcript and a 7 kb transcript). This analysis also revealed that human CARD-6 is expressed in the brain, spleen, lung, liver, muscle, and kidney.

25   Example 16:       CARD-6 increases intracellular signaling.

          The studies described in this Example demonstrate that CARD-6 expression can increase intracellular signaling.

          In a first study, a vector which expresses rat CARD-6 under the control of a CMV promoter was transiently transfected into 293 cells along with  
30 pNF $\kappa$ B-Luc (Stratagene Inc., LaJolla, CA). The pNF $\kappa$ B-Luc vector is a reporter

plasmid in which a luciferase gene is under the control of a promoter which includes a TATA box and five NF $\kappa$ B binding elements. Cotransfection of the rat CARD-6 expression vector increased luciferase expression by pNF $\kappa$ B-Luc 18-fold over that observed in the absence of the rat CARD-6 expression vector. This result indicates that CARD-6 stimulates a signaling pathway involving NF- $\kappa$ B.

In a second study, a vector expressing CARD-6 under the control of the CMV promoter was transiently transfected into 293 cells along with pAP-1-Luc (Stratagene, Inc.). The pAP-1-Luc vector is a reported plasmid in which a luciferase gene is under the control of a promoter which includes a TATA box and seven AP-1 binding sites. Co-transfection of the rat CARD-6 expression vector increased luciferase expression by pAP-1-Luc 4-fold over that observed in the absence of the rat CARD-6 expression vector. This result indicates that CARD-6 stimulates a signaling pathway involving AP-1.

Additional studies suggest that CARD-6 can stimulate phosphorylation of CHOP (GADD153), possibly by activating the stress activated kinase, JNK/p38.

#### Example 17: Deposit of Clones.

A plasmid containing a cDNA encoding human CARD-3 (pXE17A) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on May 14, 1998, and assigned Accession Number 203037.

A plasmid containing a cDNA encoding human CARD-4L (pC4L1) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on July 7, 1998, and assigned Accession Number 203035.

A plasmid containing a cDNA encoding human CARD-4S (pDB33E) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on May 14, 1998, and assigned Accession Number 203036.

A plasmid containing a cDNA encoding murine CARD-5 (EpMC5) was deposited with the American Type Culture Collection (ATCC), Manassas, VA on June 11, 1999, and assigned Accession Number PTA-212.

A plasmid containing a cDNA encoding rat CARD-6 (EpRC6) was deposited with the American Type Culture Collection (ATCC), Manassas, VA, on June 11, 1999, and assigned Accession Number PTA-211.

A clone (EpHC5) containing a cDNA molecule encoding human  
 5 CARD-5, a clone (EpCH6e) containing a cDNA molecule encoding an amino  
 terminal portion of human CARD-6, a clone (EpHC6c) containing a cDNA  
 molecule encoding a carboxy terminal portion of human CARD-6, and a clone  
 (EpHC6) containing a cDNA molecule encoding human CARD-6 were deposited  
 with the American Type Culture Collection (ATCC) Manassas, VA on June 11,  
 10 1999, as a composite deposit and assigned Accession Number PTA-213. To  
 distinguish the strains and isolate a strain harboring a particular cDNA clone, one  
 can first streak out an aliquot of the mixture to single colonies on nutrient  
 medium (e.g., LB plates) supplemented with 100 µg/ml ampicillin, grow single  
 colonies, and then extract the plasmid DNA from a selected colony using a  
 15 standard miniprep procedure. Next, one can digest a sample of the DNA  
 miniprep with a combination of the restriction enzymes Sal I and Not I and  
 resolve the resultant products on a 0.8% agarose gel using standard DNA  
 electrophoresis conditions. The digestion will liberate DNA fragments as  
 follows:

20	Human CARD-5 (EpHC5)	0.6 kb and 3.0 kb
	Human CARD-6 amino- terminal portion (EpHC6e) (amino acids 1-279)	1.0 kb and 4.3 kb
25	Human CARD-6 carboxy terminal portion (EpHC6c) (amino acid 93-1037)	3.8 kb and 3.0 kb
30	Human CARD-6 (EpHC6) (amino acids 1-1037)	4.2 kb and 3.0 kb

### Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to

5 be encompassed by the following claims.

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

- a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to the nucleotide sequence of SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, or a complement thereof;
- b) a nucleic acid molecule comprising a fragment of at least 300 nucleotides of the nucleotide sequence of SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, or a complement thereof;
- c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_;
- d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the polypeptide encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the polypeptide encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_; and
- e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, wherein the nucleic acid molecule hybridizes to a

30

nucleic acid molecule comprising SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or a complement thereof under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, which is  
5 selected from the group consisting of:
- a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, or a complement thereof; and
  - 10 b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_.
- 15 3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.
4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.
- 20 5. A host cell which contains the nucleic acid molecule of claim 1.
6. The host cell of claim 5 which is a mammalian host cell.
- 25 7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:
- a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, or 61, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, or 61;
- b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, or 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_,
- wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or a complement thereof under stringent conditions; and
- c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 65% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or a complement thereof.
9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, or 61.
10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.
11. An antibody which selectively binds to a polypeptide of claim 8.
12. A method for producing a polypeptide selected from the group consisting of:
- a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by



the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_;

5           b)     a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Number \_\_\_\_\_, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_; and

10           c)     a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 8, 26, 39, 41, 43, 49, 52, 55, 61, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as any of Accession Numbers \_\_\_\_\_, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes  
15     to a nucleic acid molecule comprising SEQ ID NO:1, 3, 7, 9, 25, 27, 38, 40, 42, 48, 50, 51, 53, 54, 56, 60, 62, or a complement thereof under stringent conditions; comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

20           13.     A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

          a)     contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

          b)     determining whether the compound binds to the polypeptide in  
25     the sample.

          14.     The method of claim 13, wherein the compound which binds to the polypeptide is an antibody.

30

15. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

- a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and
- b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

10

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

15

19. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

- a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and
- b) determining whether the polypeptide binds to the test compound.

20

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

- a) detection of binding by direct detecting of test compound/polypeptide binding;
- b) detection of binding using a competition binding assay;

30

c) detection of binding using an assay for CARD-3, CARD-4, CARD-5, or CARD-6-mediated signal transduction.

21. A method for modulating the activity of a polypeptide of claim  
5 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim  
8 with a compound which binds to the polypeptide in a sufficient concentration to  
modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the  
10 activity of a polypeptide of claim 8, comprising:  
a) contacting a polypeptide of claim 8 with a test compound; and  
b) determining the effect of the test compound on the activity of  
the polypeptide to thereby identify a compound which modulates the activity of  
the polypeptide.

15

1/58

TCACCCCTCCCGTCAGCTCTGGTTGGAGAAGCAGCGGCTGGCGTGGGGCCATCCGGGGAAATGGGC  
GCCCTCGTGACCTAGTGTTCGCGGGGCAAAAAGGGTCTTCCCGGCTCGCTCGTGCGAGGGGCTAT  
TTGGGCGCCTGAGCGCGGCGTGGGAGCCTTGGGAGCGCGCCGAGCAGGGGGGACACCGCGGAACCC  
CCCTGAGCGCGCGGGACCATGAACCGGGAGGGCCATCTGCAGCGCGCTGGCCACCATTCGCTACCA  
CAAACTCGCGGACCTCGCGTACCTGAGCCCGGGCGCTCTGGCACTGTGTCTGTCGCGCGCGCGCAG  
CAGACTGGCGCGTCCAGGTGGCCCTGAAGCACCTGCACATCCACACTCCGCTGCTGACAGTGAA  
AGAAAGGATGTCTTAAGAGAAGCTGAAATTTTACACAAAGCTAGATTAGTTACATTTCTTCAAT  
TTTGGGAATTTGCAATGAGCCTGAAATTTTGGGAATAGTTACTGAATACATGCCAATGGATCAT  
TAAATGAACCTCTACATAGGAAAACCTGAATATCCTGATGTGGCTTGGCCATTGAGATTTCTATC  
CTGCATGAAATTTGCCCTTGGTGTAAATTACCTGCACAATATGACTCCTCCTTTACTTCATCATGA  
CTTGAAGACTCAGAATATCTTATTGGACAATGAATTTCTGTTAAGATTGCAGATTTTGGTTTAT  
CAAAGTGGCGCATGATGTCCCTCTCACAGTCACGAAGTAGCAAATCTGCACCAGAAGGAGGGACA  
ATTATCTATATGCCACCTGAAAACCTATGAACCTGGACAAAAATCAGGGGCCAGTATCAAGCACGA  
TATATATAGCTATGCAGTTATCAGATGGGAAGTGTATCCAGAAAACAGCCCTTTGAAGATGTCA  
CCAATCCTTTGCAGATAATGTATAGTGTGTCAAGGACATCGACCTGTTATTAATGAAGAAAGT  
TTGCCATATGATATACCTCACCGAGCACGTATGATCTCTCTAATAGAAAGTGGATGGGCAAAAA  
TCCAGATGAAAGACCATCTTTCTTAAAATGTTTAAATAGAATTTGAACCAAGTTTGAACATTG  
AAGAGATTAATTTCTTGAAGCTTTTATTCAGCTTAAGAAAACAAAGTTACAGAGTGTTCAGT  
GCCATTACCTATGTGACAAGAAGAAAATGGAATTATCTCTGAACATACCTGTAAATCATGGTCC  
ACAAGAGGAATCATGTGGATCTCTCAGCTCCATGAAAATAGTGGTTCTCTGAAACTTCAAGGT  
CCCTGCCAGCTCCTCAAGACAATGATTTTATCTAGAAAAGCTCAAGACTGTTATTTTATGAAG  
CTGCATCACTGTCTTGGAAATCAAGTTGGGATAGCACCATTTCTGGATCTCAAAGGGCTGCATT  
CTGTGATCACAAGACCATTTCTATGCTTTTCAAGCAATAATAATCCACTCTCAACTGCAGGAACT  
CAGAACGTCTGCAGCCTGGTATAGCCGAGCAGTGGATCCAGAGCAAAAGGGAAGACATTTGTGAAC  
CAAATGACAGAAGCCTGCCCTTAACCAAGTCGCTAGATGCCCTTCTGTCCAGGGACTTATCATGAA  
AGAGGACTATGAACTTGTAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACA  
CTACTGACATCCAAGGAGAAGAATTTGCCAAGTTATAGTACAAAAATTGAAAGATAACAAACAA  
ATGGGTCTTTCAGCCTTACCCGGAATACTTGTGGTTTCTAGATCACCATCTTTAAATTTACTTCA  
AAATAAAAGCATGTAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAAAGGATATTTATAAA  
AA (SEQ ID NO:1)

FIG. 1

2/58

Met	Asn	Gly	Glu	Ala	Ile	Cys	Ser	Ala	Leu	Pro	Thr	Ile	Pro	Tyr	His	Lys	Leu	Ala	Asp	20
Leu	Arg	Tyr	Leu	Ser	Arg	Gly	Ala	Ser	Gly	Thr	Val	Ser	Ser	Ala	Arg	His	Ala	Asp	40	
Arg	Val	Gln	Val	Ala	Val	Lys	His	Leu	His	Ile	His	Thr	Pro	Leu	Leu	Asp	Ser	Gln	Arg	60
Lys	Asp	Val	Leu	Arg	Glu	Ala	Glu	Ile	Leu	His	Lys	Ala	Arg	Phe	Ser	Tyr	Ile	Leu	Pro	80
Ile	Leu	Gly	Ile	Cys	Asn	Glu	Pro	Glu	Phe	Leu	Gly	Ile	Val	Thr	Glu	Tyr	Met	Pro	Asn	100
Gly	Ser	Leu	Asn	Glu	Leu	Leu	His	Arg	Lys	Thr	Glu	Tyr	Pro	Asp	Val	Ala	Trp	Pro	Leu	120
Arg	Phe	Arg	Ile	Leu	His	Glu	Ile	Ala	Leu	Gly	Val	Asn	Tyr	Leu	His	Asn	Met	Thr	Pro	140
Pro	Leu	Leu	His	His	Asp	Leu	Lys	Thr	Gln	Asn	Ile	Leu	Leu	Asp	Asn	Glu	Phe	His	Val	160
Lys	Ile	Ala	Asp	Phe	Gly	Leu	Ser	Lys	Trp	Arg	Met	Met	Ser	Leu	Ser	Gln	Ser	Arg	Ser	180
Ser	Lys	Ser	Ala	Pro	Glu	Gly	Gly	Thr	Ile	Ile	Tyr	Met	Pro	Pro	Glu	Asn	Tyr	Gln	Pro	200
Gly	Gln	Lys	Ser	Arg	Ala	Ser	Ile	Lys	His	Asp	Ile	Tyr	Ser	Tyr	Ala	Val	Ile	Thr	Trp	220
Ala	Val	Leu	Ser	Arg	Lys	Gln	Pro	Phe	Glu	Asp	Val	Thr	Asn	Pro	Leu	Gln	Ile	Met	Tyr	240
Ser	Val	Ser	Gln	Gly	His	Arg	Pro	Val	Ile	Asn	Glu	Glu	Ser	Leu	Pro	Tyr	Asp	Ile	Pro	260
His	Arg	Ala	Arg	Met	Ile	Ser	Leu	Ile	Glu	Ser	Gly	Trp	Ala	Gln	Asn	Pro	Asp	Gln	Arg	280
Pro	Ser	Phe	Leu	Lys	Cys	Leu	Ile	Glu	Leu	Glu	Pro	Val	Leu	Arg	Thr	Phe	Glu	Gln	Ile	300
Thr	Phe	Leu	Glu	Ala	Val	Ile	Gln	Leu	Lys	Lys	Thr	Lys	Leu	Gln	Ser	Val	Ser	Ser	Ala	320
Ile	His	Leu	Cys	Asp	Lys	Lys	Lys	Met	Glu	Leu	Ser	Leu	Asn	Ile	Pro	Val	Asn	His	Gly	340
Pro	Gln	Glu	Glu	Ser	Cys	Gly	Ser	Ser	Gln	Leu	His	Glu	Asn	Ser	Gly	Ser	Pro	Gln	Thr	360
Ser	Arg	Ser	Leu	Pro	Ala	Pro	Gln	Asp	Asn	Asp	Phe	Leu	Ser	Arg	Lys	Ala	Gln	Asp	Cys	380
Tyr	Phe	Met	Lys	Leu	His	His	Cys	Pro	Gly	Asn	His	Ser	Trp	Asp	Ser	Thr	Ile	Ser	Gly	400
Ser	Gln	Arg	Ala	Ala	Phe	Cys	Asp	His	Lys	Thr	Ile	Pro	Cys	Ser	Ser	Ala	Ile	Ile	Asn	420
Pro	Leu	Ser	Thr	Ala	Gly	Asn	Ser	Glu	Arg	Leu	Gln	Pro	Gly	Ile	Ala	Gln	Gln	Trp	Ile	440
Gln	Ser	Lys	Arg	Glu	Asp	Ile	Val	Asn	Gln	Met	Thr	Glu	Ala	Cys	Leu	Asn	Gln	Ser	Leu	460
Asp	Ala	Leu	Leu	Ser	Arg	Asp	Leu	Ile	Met	Lys	Glu	Asp	Tyr	Glu	Leu	Val	Ser	Thr	Lys	480
Pro	Thr	Arg	Thr	Ser	Lys	Val	Arg	Gln	Leu	Leu	Asp	Thr	Thr	Asp	Ile	Gln	Gly	Glu	Glu	500
Phe	Ala	Lys	Val	Ile	Val	Gln	Lys	Leu	Lys	Asp	Asn	Lys	Gln	Met	Gly	Leu	Gln	Pro	Tyr	520
Pro	Glu	Ile	Leu	Val	Val	Ser	Arg	Ser	Pro	Ser	Leu	Asn	Leu	Leu	Gln	Asn	Lys	Ser	Met	540

(SEQ ID NO:2)

FIG. 2

3/58

TTTTATGGG AATCGCAGCT TGAAGAGAC AGARCAATTC CAGAAWTAAA TTTRAATTCA  
AGATTTRACC AATGTTGTTT TAAATATTC TAACTTCAAA GAATGATGCC AGAACTTAA  
AAGGRCCTGC GCAGAGTAGC AGGGGCCCTG GAGGGCGCGG CCTGAATCCT GATTGCCCTT  
TTCTGAGAG GACACACGCA GCTGAAGATG AATTGGGA AAGTAGCCGC TTGCTACTTT  
AACTATGGAA GAGCAGGGCC ACAGTGAGAT GGAATAATC CCATCAGAGT CTCACCCCA  
CATTCAATTA CTGAAAAGCA ATCGGGAAC TGTGTCAC TACATCCCA ATACTCAGTG  
TCTGGTGGAC AACTTCTGA AGAATGACTA CTTCTCGGCC GAAGATGCGG AGATTGTGTG  
TCCCTGCCCC ACCCAGCCTG ACAAGGTCCG CAAATTCCTG GACCTGTAC AGAGCAAGGG  
CGAGGAGGTG TCCGAGTTCT TCTCTACTT GCTCCAGCAA CTCGCAGATG CCTACGTGGA  
CCTCAGGCTT TGGCTGCTGG AGATCGGCTT CTTCCCTTCC CTGCTCACTC AGAGCAAAGT  
CGTGGTCAAC ACTGACCCAG TGAGCAGGTA TACCCAGCAG CTGCGACACC ATCTGGGCCG  
TGACTCRAAG TTGCTGCTGT GCTATGCCCA GAAGGAGGAG CTGCTGCTGG AGGAGATCA  
CATGGACACC ATCATGGAGC TGGTTGGCTT CAGCAATGAG AGCCTGGCA GCTGAACAG  
CCTGGCCTGC CTCCTGGACC ACACCACCCG CATCCTCAAT GAGCAGGGTG AGACCATCTT  
CATCCTGGGT GATGCTGGG TGGCAAGTC CATGCTGCTA CAGCGGCTGC AGAGCCTCTG  
GGCCACGGGC CGCTAGACG CAGGGGTCAA ATTCTCTTC CACTTTCCTT GCCGCATGTT  
CAGCTGCTTC AAGGAAAGTG ACAGGCTGTG TCTGCAGGAC CTGCTCTTCA AGCACTACTG  
CTACCCAGAG CGGGACCCCG AGGAGGTGTT TSCCTTCCTG CTGCGCTTC CCCACGTGGC  
CCTCTTCACC TTGATGGCC TGGACGAGCT GCACTCGGAC TTGGACCTGA GCCGCGTGGC  
TGACAGCTCC TSCCCTGGG AGCCTGCCC CCCCCTGGTC TTGCTGGCCA ACCTGCTCAG  
TGGGAGCTG CTCAGGGGG CTAGCAAGCT GTCACAGCC CSCACAGGCA TTAGGTGCG  
GCGCCAGTTC CTGCGGAAGA AGGTGCTTCT CCGGGCTTC TCCCCAGCC ACCTGCGCGC  
CTATGCCAGG AGGATGTTCC CCGAGCGGC CCTGCAGGAC CGCTGCTGA GCCAGCTGA  
GGCCAACCCC AACCTCTGCA GCCTGTGCTC TGTGCCCTC TTCTGCTGGA TCATCTTCG  
GTGCTTCAG CACTTCCGTG CTGCTTTGA AGGCTCACC CAGCTGCCC ACTGCACGAT  
GACCTTGACA GATGCTTCC TCTGGTCAC TGAGGTCCAT CTGAACAGGA TSCAGCCAG  
CAGCCTGGTG CAGCGGAACA CAGCAGCCC AGTGGAGACC CTCACGCCG GCCGGGACAC  
TCTGTGCTG CTGGGGCAGG TGGCCACCG GGGCATGGAG AAGAGCCTCT TTGCTTTCAC  
CCAGGAGGAG GTGCAGGCT CCGGGCTGCA GGAGAGAGAC ATGCAGCTGG GCTTCTGCG  
GGCTTTCGG GAGCTGGGCC CCGGGGTGA CAGCAGTCC TATGAGTTT TCCACCTCAG  
CCTCAGGCC TTCTTACAG CTTCTTCTT CTTGCTGGAC GACAGGGTGG GCACTCAGGA  
GCTGCTCAGG TTCTTCAGG AGTGATGCC CCTGCGGGG GCAGCGACA CTTCTGCTA

4/58

TTTTCCCTTC CTCCCGTTCC AGTGCCCTCA GGGCAGTGGT CCGCGCGGGG AAGACCTCTT  
TAAGAACAAG GATCACTTCC AGTTCACCAA CCTCTTCCTG TCGGGGTGT TGTCCAAAGC  
TAAACAGAAA CTCTGCGGC ATCTGCTGCC CCGGGCAGCC CTGAGGAGAA AGCGCAAGGC  
CTGTGGGGCA CACCTGTTTT CCAGCCTCCG GGGCTACCTG AAGAGCCTGC CCCGCGTTCA  
CTCGAAAGC TTCAACCAGG TCCAGGCCAT CCCCACGTTT ATCTGGATGC TCGGCTGCAT  
TTACGAGACA CAGAGCCAGA AGGTGGGGCA GCTGGCGGGC AGGGGCATCT GCGCCAACTA  
CTCAAGCTG ACCTACTGCA ACGCCTCTC GCGCGACTGC AGCGCCCTCT CCTTCCTCTT  
GCATCACTTC CCCAAGCGGC TGGCCCTAGA CCTAGACAA CACAATCTCA ACGACTACGG  
CGTGGGGGAG CTGCAGCCCT GCTTCAGCCG COTCACTGTT CTCAGACTCA GCGTAAACCA  
GATCACTGAC GGTGGGGTAA AGGTGCTAAG CGAAGAGCTG ACCAAATACA AAATTGTGAC  
CTATTGGGT TTATACAACA ACCAGATCAC CGATGTCCGA GCCAGGTACG TCACCAAAAT  
CTGTGATGAA TSCAAAGGCC TCAGGCATCT TAAACTGGGA AAAACAAAA TAACAAGTGA  
AGGAGGGGAG TATCTGCCCC TGGCTGTGAA GAACAGCAA TCAATCTCTG AGGTTGGGAT  
GTGGGGCAAT CAAGTTGGGG ATGAAGGAGC AAAAGCCTTC GCAGAGGCTC TCGGGAACCA  
CCCCAGCTTG ACCACCTTGA GTCTTSCGTC CAACGGCATC TCCACAGAAG GAGGAAAGAG  
CCTTGCGAGG GCCCTGCAGC AGAACACGTC TCTAGAAATA CTGTGGCTGA CCCAAAATGA  
ACTCAACGAT GAAGTGGCAG AGAGTTTGGC AGAAATGTTG AAAGTCAACC AGACGTTAAA  
GCATTATGG CTATCCAGA ATCAGATCAC AGCTAAGGGG ACTGCCCAGC TGGCAGATGC  
GTTACAGAGC AACACTGGCA TAACAGAGAT TTSCCTAAAT GGAAACCTGA TAAAACCAGA  
GGAGGCCAAA GTCTATGAAG ATGAGAAGCG GATTATCTGT TTCTGAGAGG ATGCTTTCTT  
GTCATGGGG TTTTGGCCCT GGAGCCTTAG CAGCAAATGC CACTCTGGGC AGTCTTTTGT  
GTCAGTGTCT TAAAGGGGGC TCGGCAGGCG GGACTATCAG GAGTCCACTG CCTYCATGAT  
GCAAGCCAGC TTCTGTGCA GAAGGTCTGG TCGGCAAACT CCTAAGTAC CCGCTACAA  
TCTGCAGAAA AAGAATGTGT CTTGCGAGCT GTTGTAGTGA CAGTAAATAC ACTGTGAAGA  
GAAAAAAAAA ACGGACGCGT GG (SEQ ID NO:7)

FIG. 3 (page 2 of 2)

5/58

MEEQCHSEMEIIPSESHPHIQLLKSNRELLVTHIRNTQCLVDNLLKNDYFSAEDAEIYCACPTQP  
DKVRKILDVQSKGEEVSEFFLYLLQQLADAYVDLRPWLEIGFSPSLLTQSKVVVNTDPVSRYT  
QQLRHHLGRDSKFVLCYAQKEELLLEETVMDTIMELVGFSNESLGSLNSLACLLDHTTGILNEQG  
ETIFILGDAGVGKSMILLQRLQSLWATGRLEAGVKFFHFRCRMFSCFKESDRLCLQDLLFKHYCY  
PERDPHEVFAFLLRFPHVALFTFDGLDELHSDLDLSRVPDSSCPWEPAPLVLNLLSGKLLKG  
ASKLLTARTGIEVPRQFLRKKVLLRGFSPSHLRAYARRMFPERALQDRLLSQLEANPNLCSLCSV  
PLFCWIIIFRCFQHFRAAFEGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNTRSPVETLHA  
GRDTLCSLGQVAHRGMEKSLFVFTQEEVQASGLQERDMQLGFLRALPELGPGGDQQSYYEFFHLTL  
QAFFTAFFFLVDDRVTGTCELLRFFQEWMPAGAAATTSCYPFFLFFQCLQGSGPAREDLFANKDHF  
QFTNLFLCGLLSKAKQKLLRHLVPAAALRRKRKALWAHLFSSLRGYLKSLPRVQVESFNQVQAMP  
TFIWMRLRCIYETQSQKVGQLAARGICANYLKLTYCNACSADECSALSFVLHHPKRLALDLNNDNL  
NDYGVRELQPCFSRLTVLRLSVNQITDGGVKVLSEELTKYKIVTYLGLYNNQITDVGARYVTKIL  
DECKGLTHLKLKGNKITSEGGKYLALAVKNSKSISEVGMWGNQVGDEGAKAFAEALRNHPSLTTL  
SLASNGISTEGGKSLARALQQNTSLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLIQNQITA  
KGTAQLADALQSNTGITEICLNGNLIKPEEAKVYEDEKRIICF (SEQ ID NO:8)

FIG. 4



6/58

TACGGCTCCSACTTCTGGAAGATGACTACTTCTCGGCCGAAGATGCGGAGATTGTGT  
BTGCCTGCCCCACCCAGCCTGACAAGGTCCGCAAAATTCTGGACCTGGTACAGAGCAAG  
GGCGAGGAGGTGTCCGAGTTCTTCTCTACTTGTCTCCAGCAACTCGCAGATGCCTACGT  
GGACCTCAGGCCTTGGCTGTGGAGATCGGCTTCTCCCCCTTCCCTGCTCACTCAGAGCA  
AAGTCTGGTCAACACTGACCCAGTGAGCAGGTATACCCAGCAGCTGCGACACCATCTG  
GGCCGTGACTCCAAGTTCGTGTGTGCTATGCCAGAAGGAGGAGCTGCTGCTGGAGGA  
GATCTACATGGACACCATCATGGAGCTGGTTGGCTTCAGCAATGAGAGCCTGGGCAGCC  
TGAACAGCCTGGCCTGCCTCTGGACCAACACCACCGGCATCCTCAATGAGCAGGGTGAG  
ACCATCTTCATCCTGGGTGATGCTGGGGTGGGCAAGTCCATGCTGCTACAGCGGCTGCA  
GAGCCTCTGGGCCACGGGCGGGCTAGACGCAGGGGTCAAATTCCTTCTTCCACTTTCGCT  
GGCGCATGTTCAGCTGCTTCAAGGAAAGTGACAGGCTGTGTCTGCAGGACCTGCTCTTC  
AAGCACTACTGCTACCCAGAGCGGGACCCCGAGGAGGTGTTTGCCTTCTGCTGCGCTT  
CCCCACGTGGCCCTTTCACCTTCGATGGCCTGGACGAGCTGCACTCGGACTTGGACC  
TGAGCCCGTGCCTGACAGCTCTGCCCCCTGGGAGCCTGCCCCACCCCTGGTCTTGCTG  
GCCAACCTGCTCAGTGGGAAGCTGCTCAAGGGGGCTAGCAAGCTGCTCACAGCCCGCAC  
AGGCATCGAGGTCCCGCGCCAGTTCTGCGGAAGAAGGTGCTTCTCCGGGGCTTCTCCC  
CCAGCCACCTGCGCGCCTATGCCAGGAGGATGTTCCCCGAGCGGGCCCTGCAGGACCGC  
TTGCTGAGCCAGCTGGAGGCCAACCACAACCTCTGCAGCCTGTGCTCTGTGCCCTCTT  
CTGCTGGATCATCTTCCGGTGTCTCCAGCACTTCCGTGCTGCCTTTGAAGGCTCACCAC  
AGCTGCCCCACTGCACGATGACCTTGACAGATGTCTTCTCCTGGTCACTGAGGTCCAT  
CTGAACAGGATGCAGCCCAGCAGCCTTGTGTCAGCGGAACACACGCAGCCCAGTGGAGAC  
CCTCCACGCGCGCGGACACTCTGTGCTCGCTGGGGCAGGTGGCCCCACCGGGGCATGG  
AGAAGAGCCTCTTTGTCTTACCAGGAGGAGGTGCAGGCCTCCGGGCTGCAGGAGAGA  
GACATGCAGCTGGGCTTCTGCGGGCTTTGCCGGAGCTGGGCCCCGGGGGTGACCAGCA  
GTCTTATGAGTTTTTCCACCTCAGCCTCCTCACCTGTAAACTGGGATCCCAGTATAGA  
CTTTGGAAATCAGTAGACACCATATGCTTCAAAAAACAGGGGCTATTAAATGACATCA  
GGAGCCAGAAAGTCTCATGGCTGTGCTTTCTCTTGAAGTTTATACAACAACCAGATCAC  
CGATGTCGGAGCCAGACTGGGAAAAAACAATAAAGTGAAGGAGGGAAGTATCTCG  
CCCTGGCTGTGAAGAACAGCAAATCAATCTCTGAGGTGGGATGTGGGGCAATCAAGTT  
GGGGATGAAGGAGCAAAAGCCTTCGCAGAGGCTCTGCGGAACCAACCCAGCTTGACCAC  
CCTGAGTCTTGGCTCCAACGGCATCTCCACAGAAGGAGGAAAGAGCCTTGGCAGGGCCC  
TGCAGCAGAACACGTCTCTAGAAATACTGTGGCTGACCCAAAATGAACTCAACGATGAA  
BTGGCAGAGAGTTTGGCAGAAATGTTGAAAGTCAACCAGACGTTAAAGCATTTATGGCT  
TATCCAGAATCAGATCACAGTCTTTTGTGTGAGTGTCTTAAAGGGGCTGCGCAGGGCG  
GACTATCAGGAGTCCACTGCCTCCATGATGCAAGCCAGCTTCTGTGCAGAAGGTCTGG  
TCGGCAAACTCCCTAAGTACCGCTACAATTCTGCAGAAAAAGAATGTGPTTGGCAGC  
TGTGTAGTTACAGTAAATACACTGTGAAGAGACTTTATTGCCTATTATAATTATTTT  
ATCTGAAGCTAGAGGAATAAAGCTGTGAGCAACAGAGGAGGCCAGCCTCACCTCATT  
CAACACCTGCCATAGGGACCAACGGGAGCGAGTTGGTCACCGCTCTTTTCATTGAAGAG  
TTGAGGATGTGGCACAAGTTGGTGCCAAGCTTCTGAATAAAACGTGTTTGATGGATT  
AGTATTATACCTGAAATATTTTCTTCTTCTCAGCACTTCCCATGTATTGATACTGGT  
CCCACTTCACAGCTGGAGACACCGGAGTATGTGCAGTGTGGGATTTGACTCCTCCAAGG  
TTTTGTGAAAAGTTAATGTCAAGGAAAGGATGCACCACGGGCTTTTAATTTTAACTCTG  
GAGTCTCACTGTCTGCTGGCAAGATAGAGAATGCCCTCAGCTCTTAGCTGGTCTAAGA  
ATGACGATGCCTTCAAAATGCTGCTTCCACTCAGGGCTTCTCCTCTGCTAGGCTACCCT  
CCTCTAGAAGGCTGAGTACCATGGGCTACAGTGTCTGGCCTTGGGAAGAAGTGATTCTG  
TCCCTCCAAAGAAATAGGGCATGGCTTGGCCCTGTGGCCCTGGCATCCAAATGGCTGCT  
TTTGTCTCCCTTACCTCGTGAAGAGGGGAAGTCTTCTTCTGCTCCCAAGCAGCTGAAG  
GGTGAATAAACGGGCGCCAAGACTCAGGGGATCGGCTGGGAACCTGGGCCAGCAGAGCAT  
BTGGACACCCCCACCATGGTGGGCTTGTGGTGGCTGCTCCATGAGGGTGGGGGTGAT  
ACTACTAGATCACTTGTCTCTTCCAGCTCATTTGTTAATAAAATACTGAAAAACAA  
AA  
AAAAAAAAAAAAAA (SEQ ID NO:25)

FIG. 5

7/58

HASBLLKNDYFSAEDAEIWCACFTQPDKVRKILDLVQSKGEEVSEFFLYLL  
QQLADAYVDLRPWLLLEIGFSPSLLTQSKVVVNTDPVSRYTQQLRHHLGRDS  
KFVLCYAKKEELLLEEIYMDTIMELVGFSNESLGSLSLACLLDHTTGILN  
EQGETIFILGDAGVGKSMLLQRLQSLWATGRLDAGVKFFHFRCRMFSCFK  
ESDRICLQDILLEKHYCYPERPPEEVFAFLLRFPHVALFTFDGLDELHSDLD  
LSRVPDSSCPWEPANPLVLLANLLSGKLLKAGASKLLTARTGIEVPRQFLRK  
KVLLRGFSPSHLRAYARRMFPERALQDRLLSQLEANPNLCSLCSVPLFCWI  
IFRCFQHFRAAFEGSPQLPDCMTLTDVFLLVTEVHLNRMQPSSLVQRNTR  
SPVETLHAGRDTLCSLGQVAHRGMEKSLFVFTQEEVQASGLQERDMQLGFL  
RALPELGPGGDQOSYEFFHLSLLTCKTGIFV (SEQ ID NO:26)

FIG. 6

8/58

	10	20	30	40					
1	ESRTHC	--LLK	HEFL	LL	THIRNTQ	I--V	NL	NDY	CARD4-CARD
2	L-QG	ACW	QX	ED	IN	MEA	LMQS	LRDL	CARD3-CARD
3	AGER	--PSET	DRE	KRL	ET	QADS	G--L	AR	ARC-CARD
4	WAS	DES	--LI	KHE	MA	F--	--V	NL	CARD1-CARD
5	RES	NDIL	--EIK	HE	MA	FHL	--VI	SL	CARD2-CARD
	50	60	70	80					
35	FSAE	DASE	IC	CE	QED	K	IL	IL	CARD4-CARD
36	IMH	EDY	LE	ST	RT	S	K	VE	CARD3-CARD
37	LTG	PEY	ALD	LE	DA	ERR	RL	LL	ARC-CARD
38	YTP	Q--	DE	I	QI	TQ	I	AE	CARD1-CARD
39	IE	EQ	--	DVI	Q	TQ	TS	LA	CARD2-CARD
	90	100							
75	LE	QLE	DE	YVD	I	R	E	W	CARD4-CARD
76	IK	DN	KQ	--	MG	Q	PE	I	CARD3-CARD
77	AR	TE	GE	PD	P	A	W	D	ARC-CARD
78	LK	E	ID	ST	--	--	K	N	CARD1-CARD
79	LE	Q	AE	AV	--	--	E	H	CARD2-CARD

FIG. 7

9/58

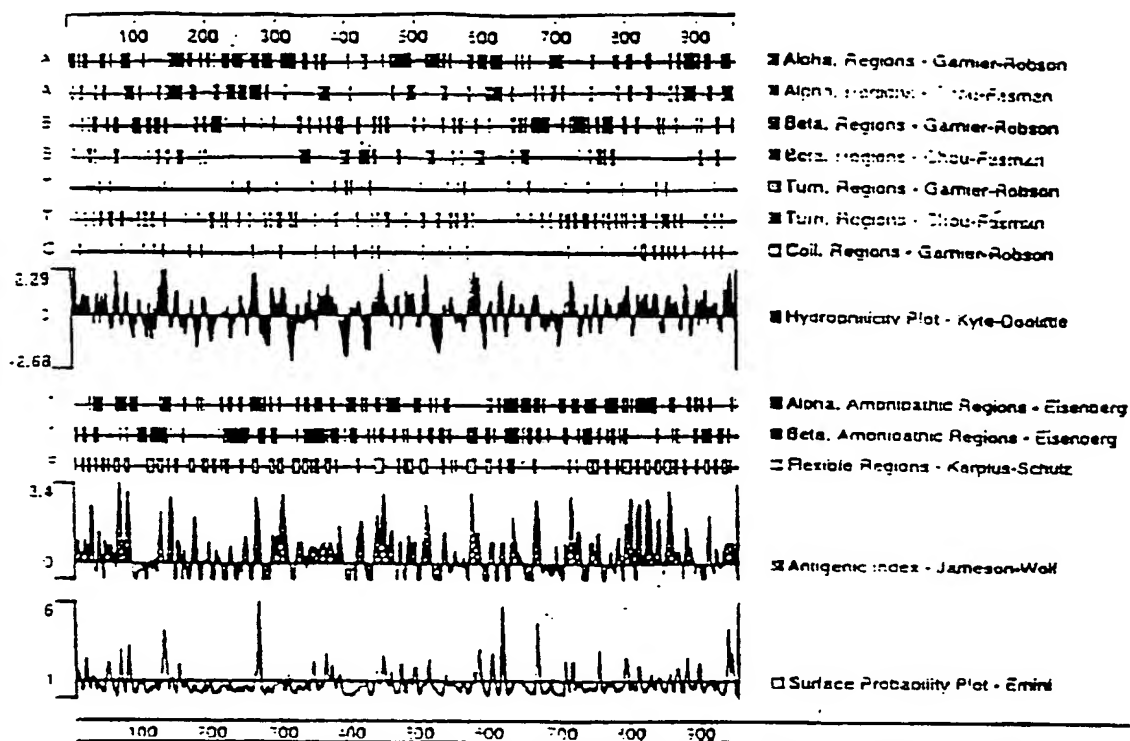


FIG. 3

10/58

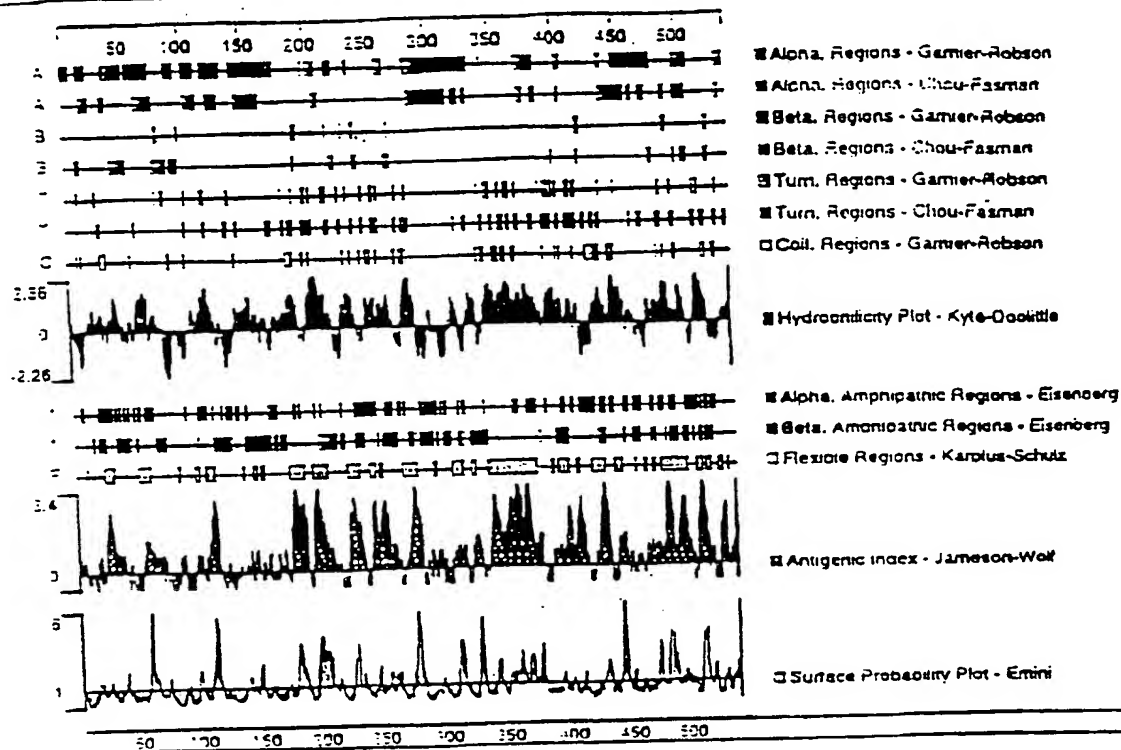


FIG. 9

11/58

CCCCGCTCCGCGTCCCCGGACCATGGCGCTCTCCGGGCTCTTCTCTAGCTCTCAGCGGCT  
GCGAAGTCTGTNAACCTGGTGGCCAAGTGATTGTAAGTCAGGAGACTTTCTTCGGTTTC  
TGCCTTTGATGGCAAGAGGTGGAGATTGTGGCGGCGATTACAGAAAACATCTGGGAAGAC  
AAGTTGCTGTTTTTATGGGAATCGCAGGCTTGGAAGAGACAGAAGCAATTCCAGAAATAA  
ATTGGAAATTGAAGATTTAAACAATGTTGTTTTAAAAATATTCTAACTTCAAAGAATGATG  
CCAGAAACTTAAAAAGGGGCTGCGCAGAGTAGCAGGGGCCCTGGAGGGCGCGGCTGAAT  
CCTGATTGCCCTTCTGCTGAGAGGACACACGCAGCTGAAGATGAATTTGGGAAAAGTAGC  
CGCTTGCTACTTTAACTATGGAAGAGCAGGGCCACAGTGAGATGGAAATAATCCCATCAG  
AGTCTCACCCCCACATTCAATTACTGAAAAGCAATCGGGAACCTTCTGGTCACTCACATCC  
GCAATACTCAGTGTCTGGTGGACAACTTGCTGAAGAATGACTACTTCTCGGCCGAAGATG  
CGGAGATTGTGTGTGCCTGCCCCACCCAGCCTGACAAGGTCCGCAAAATTCTGGACCTGG  
TACAGAGCAAGGGCGAGGAGGTGTCCGAGTTCTTCTCTACTTGCTCCAGCAACTCGCAG  
ATGCCTACGTGGACCTCAGGCCTTGGCTGCTGGAGATCGGCTTCTCCCCTTCCCTGCTCA  
CTCAGAGCAAAGTCGTGGTCAACACTGACCCAGTGAGCAGGTATACCCAGCAGCTGCGAC  
ACCATCTGGGGCCGTGACTCCAAGTTCGTGCTGTGCTATGCCCAGAAGGAGGAGCTGCTGC  
TGGAGGAGATCTACATGGACACCATCATGGAGCTGGTTGGCTTCAGCAATGAGAGCCTGG  
GCAGCCTGAACAGCCTGGCCTGCCTCCTGGACCACACCACCGGCATCCTCAATGAGCAGG  
CTGCTTCAAGGAAAGTGACAGGCTGTGTCTGCAGGACCTGCTCTTCAAGCACTACTGCTA  
CCCAGAGCGGGACCCCGAGGAGGTGTTTGCCTTCTGCTGCGCTTCCCCACGTGGCCCT  
CTTCACCTTCGATGGCCTGGACGAGCTGCACTCGGACTTGGACCTGAGCCGCGTGCCTGA  
CAGCTCCTGCCCCCTGGGAGCCTGCCCCACCCCTGGTCTTGTGCTGGCCAACCTGCTCAGTGG  
GAAGCTGCTCAAGG3GGCTAGCAAGCTGCTCACAGCCCCGACAGGCATCGAGGTCCCGCG  
CCAGTTCCTGCGGAAGAAGGTGCTTCTCCGGGGCTTCTCCCCCAGCCACCTGCGCGCCTA  
TGCCAGGAGGATGTTCCCCGAGCGGGCCCTGCAGGACCGCCTGCTGAGCCAGCTGGAGGC  
CAACCCCAACCTCTGCAGCCTGTGCTCTGTGCCCCCTCTTCTGCTGGATCATCTTCCGGTG  
CTTCCAGCACTTCCGTGCTGCCTTTGAAGGCTCACCACAGCTGCCCCGACTGCACGATGAC  
CCTGACAGATGTCTTCTCCTGGTCACTGAGGTCCATCTGAACAGGATGCAGCCCAGCAG  
CCTGGTGCAGCGGAACACACGCAGCCAGTGGAGACCTCCACGCCGCGCGGACACTCT  
GTGCTCGCTGGGGCAGGTGGCCACCGGGGCATGGAGAAGAGCCTCTTTGTCTTACCCCA  
GGAGGAGGTGCAGGCCTCCGGGCTGCAGGAGAGAGACATGCAGCTGGGCTTCTGCGGGC  
TTTGCCGGAGCTGGGCCCCGGGGTGACCAGCAGTCTATGAGTTTTTCCACCTCACCT

FIG. 10 (Page 1 of 3)

12 / 58

CCAGGCCTTCTTTACAGCCTTCTTCTCTGTGCTGGACGACAGGGTGGGCACTCAGGAGCT  
GCTCAGGTTCTTCCAGGAGTGGATGCCCCCTGCGGGGGCAGCGACCACGTCCTGCTATCC  
TCCCTTCTTCCCGTTCCAGTGCCCTGCAGGGCAGTGGTCCGGCGCGGGAAGACCTCTTCAA  
GAACAAGGATCACTTCCAGTTCACCAACCTCTTCTGTGCGGGCTGTTGKCCAAAGCCAA  
ACAGAACTCCTGCGGCATCTGGTGCCCCGCGGCAGCCCTGAGGAGAAAGCGCAAGGCCCT  
GTGGGCACACCTGTTTTCCAGCCTGCGGGGCTACCTGAAGAGCCTGCCCCGCGTTCAGGT  
CGAAAGCTTCAACCAGGTGCAGGCCATGCCCACGTTTCTGATGCTGCGCTGCATCTA  
CGAGACACAGAGCCAGAAGGTGGGGCAGCTGGCGGCCAGGGGCATCTGCGCCAACCTACCT  
CAAGCTGACCTACTGCAACGCCTGCTCGGCCGACTGCAGCGCCCTCTCCTTCGTCTGCA  
TCACTTCCCCAAGCGGCTGGCCCTAGACCTAGACAACAACAATCTCAACGACTACGGCGT  
GCGGGAGCTGCAGCCCTGCTTCAGCCGCCTCACTGTTCTCAGACTCAGCGTAAACCAGAT  
CACTGACGGTGGGGTAAAGGTGCTAAGCGAAGAGCTGACCAAATACAAAATTGTGACCTA  
TTTGGGTTTATACAACAACCAGATCACCGATGTCTGGAGCCAGGTACGTCACCAAATCCT  
GGATGAATGCAAAGGCCTCACGCATCTTAACTGGGAAAAAACAAAATAACAAGTGAAGG  
AGGGAAGTATCTCGCCCTGGCTGTGAAGAACAGCAAATCAATCTCTGAGGTTGGGATGTG  
GGGCAATCAAGTTGGGGATGAAGGAGCAAAAGCCTTCGCAGAGGCTCTGCGGAACCAACC  
CAGCTTGACCACCCTGAGTCTTGCGTCCAACGGCATCTCCACAGAAGGAGGAAAGAGCCT  
TGCGAGGGCCCTGCAGCAGAACACGTCTCTAGAAATACTGTGGCTGACCCAAAATGAACT  
CAACGATGAAGTGGCAGAGAGTTTGGCAGAAATGTTGAAAGTCAACCAGACGTTAAAGCA  
TTTATGGCTTATCCAGAATCASATCACAGCTWARGGGACTGCCAGCTGGCAGATGCGTT  
ACAGAGCAACACTGGCATAACAGAGATTTGCCTAAATGGAAACCTGATAAAACCAGAGGA  
GGCCAAAGTCTATGAAGATGAGAAGCGGATTATCTGTTTCTGAGAGGATGCTTTCCTGTT  
CATGGGGTTTTTGCCCTGGAGCCTCAGCAGCAAATGCCACTYTGCGCAGTCTTTTGTGTC  
AGTGTCTTAAAGGGGCTGCGCAGGCGGGACTATCAGGAGTCCACTGGCTCCATGATGCA  
AGCCAGCTTCTGTGCAGAAGGTCTGGTTCGGCAAACCTCCCTAAGTACCCGCTACAATTCT  
GCAGAAAAAGAATGTGTCTTGCGAGCTGTTGTAGTTACAGTAAATACACTGTGAAGAGAC  
TTTATTGCCTATTATAATTATTTTATCTGAAGCTAGAGGAATAAAGCTGTGAGCAAACA  
GAGGAGGCCAGCCTCACCTCATTCCAACACCTGCCATAGGGACCAACGGGAGCGAGTTGG  
TCACCGCTCTTTTCATTGAAGAGTTGAGGATGTGGCACAAAGTTGGTGCCAAGCTTCTTG  
AATAAAACGTGTTTGTATGATTAGTATTATACCTGAAATATTTTCTTCTCTCAGCACT  
TTCCCATGTATTGATACTGGTCCCACTTCACAGCTGGAGACACCGGAGTATGTGCAGTGT  
GGGATTTGACTCCTCCAAGGTTTTGTGGAAAGTTAATGTCAAGGAAAGGATGCACCACGG

FIG. 10 (Page 2 of 3)

13/58

GCTTTTAATTTTAATCCTGGAGTCTCACTGTCTGCTGGCAAAGATAGAGAATGCCCTCAG  
CTCTTAGCTGGTCTAAGAATGACGATGCCTTCAAAATGCTGCTTCCACTCAGGGCTTCTC  
CTCTGCTAGGCTACCCTCCTCTAGAAGGCTGAGTACCATGGGCTACAGTGTCTGGCCTTG  
GGAAGAAGTGATTCTGTCCCTCCAAAGAAATAGGGCATGGCTTGCCCTGTGGCCCTGGC  
ATCCAAATGGCTGCTTTTGTCTCCCTTACCTCGTGAAGAGGGGAAGTCTCTTCCTGCCTC  
CCAAGCAGCTGAAGGGTGACTAAACGGGCGCCAAGACTCAGGGGATCGGCTGGGAACTGG  
GCCAGCAGAGCATGTTGGACACCCCCCACCATGGTGGGCTTGTGGTGGCTGCTCCATGAG  
GGTGGGGGTGATACTACTAGATCACTTGTCTCTTGCCAGCTCATTGTTAATAAAATAC  
TGAAAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGG (SEQ ID NO:38)

FIG. 10 (Page 3 of 3)



14/58

MEEQGHSEMEIIPSESHPHIQLLKSRELLVTHIRNTQCLVDNLLKNDYFSAEDAEIVCA  
CPTQPKVRKILDLVQSKGEEVSEFFLYLLQQLADAYVDLRPWLLEIGFSPSLLTQSKVV  
VNTDPVSRYTQQLRHHLGRDSKFVLCYAKEELLLEEIYMDTIMELVGFSNESLGSLSL  
ACLLDHTTGILNEQAASRKVTGCVCRITCSSSTTATQSGTPRRCLPSCCASPTWPSSPSMA  
WTSCRTWT (SEQ ID NO:39)

FIG. 11

15/58

CACGCGTCCGCGCTACTGCGGGAGCAGCGTCCTCCCGGGCCACGGCGCTTCCCGGCCCG  
GCGTCCCCGGACCATGGCGCTCTCCGGGCTCTTCTCTAGCTCTCAGCGGCTCGGAAGTCT  
GTAAACCTGGTGGCCAAGTGATTGTAAGTCAGGAGACTTTCCTTCGGTTTCTGCCTTTGA  
TGGCAAGAGGTGGAGATTGTGGCGGCGATTACAGAAAACATCTGGGAAGACAAGTTGCTG  
TTTTTATGGGAATCGCAGGCTTGGAAGAGACAGAAGCAATTCAGAAATAAATTGGAAAT  
TGAAGATTAAACAATGTTGTTTTAAATATTCTAACTTCAAAGAATGATGCCAGAACT  
TAAAAAGGGGCTGCGCAGAGTAGCAGGGGCCCTGGAGGGCGCGGCCTGAATCCTGATTGC  
CCTTCTGCTGAGAGGACACACGCAGCTGAAGATGAATTTGGGAAAAGTAGCCGCTTGCTA  
CTTTAACTATGGAAGAGCAGGGGCCACAGTGAGATGGAAATAATCCCATCAGAGTCTCACC  
CCCACATTCAATTACTGAAAAGCAATCGGGAACCTTCTGGTCACTCACATCCGCAATACTC  
AGTGTCTGGTGGACAACTTGCTGAAGAATGACTACTTCTCGGCCGAAGATGCGGAGATTG  
TGTGTGCTGCCCCACCCAGCCTGACAAGGTCCGCAAAATTCTGGACCTGGTACAGAGCA  
AGGGCGAGGAGGTGTCCGAGTTCTTCCTCTACTTGCTCCAGCAACTCGCAGATGCCTACG  
TGGACCTCAGGCCTTGCGTGCTGGAGATCGGCTTCTCCCCTTCCCTGCTCACTCAGAGCA  
AAGTCGTGGTCAACACTGACCCAGGTAGGAGTCAGCCCCAGCAAGACCGCAGGCACCACT  
GCAAGCAGGGCCCTGGGGGGTTTGGTAATGGCTGGGCCAGCCCTGAGTGCCACCTCAGGA  
AGCAGGCCCAGGTGCTATTTTGATTTTAGAAAGGAACAGCTGAATCCTGTCTCCCAAGTG  
CAGCCCAGGTGGCTGCGATTGAACTGCCCACACCTCGATGGTCTGGTTTATAGAGGGGCC  
TTTGAAGTATGGGAATGGCCTGTGTTCTGACCCCTTGCTTCTTCCTATTCTGACATAT  
GTAGACATTTTAATGGTTGCACAAATTCAGGTTGTATTTTTTTTTCTTTAAAAAAATCT  
TTAGCTGGACATGGTAGCACACACCTGTAGTTCCAGCTACTCAGGAGGCTGAGGCAAGAG  
EACTGCTTGAGCCCCAGAGTCTAAGGCTGCAGCGAGCTATGATTGTGCCCCTACACTCCA  
CAGCCTGGGTTTTAGAGTGAGACCCTGTCTCTAAAAA  
AA  
AAAAAAAAAAAAAAAAANGGGCGG (SEQ ID NO:40)

FIG. 12

16/58

MEEQGHSEMEIIPSESHPHIQLLKSNRELLVTHIRNTQCLVDNLLKNDYFSAEDAEIVCA  
CPTQPKVRKILDLVQSKGEEVSEFFLYLLQQLADAYVDLRPWLLEIGFSFSLLTQSKVV  
VNTDPGRSQPQQDRRHQCKQGFGGFGNGWASPECHLRKQAQVLF  
(SEQ ID NO:41)

FIG. 13

17/58

	MEEQGHSEMEIIPSESHPHIQLLKSNNRELLVTHIRNTQCL	Majority
	10 20 30 40	
1	MEEQGHSEMEIIPSESHPHIQLLKSNNRELLVTHIRNTQCL	CARD4-Y CLONE
1	MEEQGHSEMEIIPSESHPHIQLLKSNNRELLVTHIRNTQCL	CARD4-Z CLONE
1	MEEQGHSEMEIIPSESHPHIQLLKSNNRELLVTHIRNTQCL	CARD4L
	VDNLLKNDYFSAEDAEIVCACPTQPKVRKILDLVOSKGE	Majority
	50 60 70 80	
11	VDNLLKNDYFSAEDAEIVCACPTQPKVRKILDLVOSKGE	CARD4-Y CLONE
11	VDNLLKNDYFSAEDAEIVCACPTQPKVRKILDLVOSKGE	CARD4-Z CLONE
11	VDNLLKNDYFSAEDAEIVCACPTQPKVRKILDLVOSKGE	CARD4L
	EVSEFFLYLLOQLADAYVDLRPWLLLEIGFSPSLLTOSKVV	Majority
	90 100 110 120	
31	EVSEFFLYLLOQLADAYVDLRPWLLLEIGFSPSLLTOSKVV	CARD4-Y CLONE
31	EVSEFFLYLLOQLADAYVDLRPWLLLEIGFSPSLLTOSKVV	CARD4-Z CLONE
31	EVSEFFLYLLOQLADAYVDLRPWLLLEIGFSPSLLTOSKVV	CARD4L
	VNTDPVSRYTQQLRHHLGGRDSKFVLCYAOKEELLLEEIYM	Majority
	130 140 150 160	
121	VNTDPVSRYTQQLRHHLGGRDSKFVLCYAOKEELLLEEIYM	CARD4-Y CLONE
121	VNTDPGRSOPQC DRR - - - - -	CARD4-Z CLONE
121	VNTDPVSRYTQQLRHHLGGRDSKFVLCYAOKEELLLEEIYM	CARD4L
	DTIMELVGFSNESLGSLSLACLLDHTTGILNEOXXXX - -	Majority
	170 180 190 200	
151	DTIMELVGFSNESLGSLSLACLLDHTTGILNEOAA SR - -	CARD4-Y CLONE
137	- - - - -	CARD4-Z CLONE
151	DTIMELVGFSNESLGSLSLACLLDHTTGILNEOGETIFI	CARD4L
	- - - - - X C X X X	Majority
	210 220 230 240	
139	- - - - - V T G	CARD4-Y CLONE
137	- - - - - Q C H Q - -	CARD4-Z CLONE
231	LG DAGVGKSM L L Q R L Q S L W A T G R L D A G V K F F F H F R C R M F S	CARD4L
	- - - - - X C - - - - -	Majority
	250 260 270 280	
133	G - - - - - V G - - - - -	CARD4-Y CLONE
141	- - - - -	CARD4-Z CLONE
141	G F H E S D R L G L Q D L L F K X Y C Y P E R D P E E V F A F L L R F P H V A L	CARD4L
	- - - - - X X C X X X - - - - -	Majority
	290 300 310 320	
136	- - - - - R T G S S S - - - - -	CARD4-Y CLONE
141	- - - - -	CARD4-Z CLONE
151	F T F D G L D E L H S D L D L S R V P D S S G P W E P A H P L V L L A N L L S G	CARD4L

FIGURE 14 (1 of 4)

18/58

-----TXXTXXXXP-----Majority			
	330	340	350 360
222	-----T T A T Q E E T P R-----		
141	-----T T A T Q E E T P R-----		
321	H L L N G A S K L L T A R T G I E V P R Q F L R K K V L L R G F S P S H L R A Y		
	-----R C-----Majority		
	370	380	390 400
222	-----R C-----		
141	-----R C-----		
361	A A R M F P E R A L Q D R L L S Q L E A N P M L C S L C S V P L F C W I T F R C		
	-----L P X C-----Majority		
	410	420	430 440
224	-----L P S C-----		
141	-----L P S C-----		
401	F R H F R A A F E G S P Q L P D C T H T L T D V F L D V T E V H L N R M Q P S S		
	-----Majority		
	450	460	470 480
228	-----		
141	-----		
441	L V Q R N T R S P V E T L H A G R D T L C S L G Q V A H R G M E K S L F V F T Q		
	-----G P G G-----Majority		
	490	500	510 520
228	-----		
141	-----G P G G-----		
481	E E V Q A S G L Q E R D M Q L Q F L R A L P E L G P G G D Q Q S Y E F F H L T L		
	-----F X X X W X X P-----C X X-----Majority		
	530	540	550 560
229	-----G A S-----		
145	-----F G N G F A S P-----		
521	A A F F T A F F L V L D D R V G T Q E L L R F F Q E N M P D A G A A T T S G Y P		
	-----Majority		
	570	580	590 600
231	P-----		
153	-----		
551	P F L P F Q C L Q G S G P A R E D L F E N K D H F Q F T N L F L C G L L S K A K		
	-----X W X-----Majority		
	610	620	630 640
232	-----T W P-----		
153	-----		
591	I E L L R H L V P A A A L A R K K A L W A H L F S S L R G Y L K S L P R V Q V		

FIGURE 14 (2 of 4)

19/58

-----XXPXXXX-----				Majority
	650	660	670	680
135	-----SSRSMAR-----			CARD4-Y CLONE
153	-----SSRSMAR-----			CARD4-Z CLONE
541	EEFNQVQAMRTFTMLRCIYETQSQKVGQLAARGICANYL			CARD4L
-----XXCXXXX-----				Majority
	690	700	710	720
142	-----FSRTT-----			CARD4-Y CLONE
153	-----FSRTT-----			CARD4-Z CLONE
531	ELTYCNAQSADCSALSFVLHHPKRLALDLDNNNLNDYGV			CARD4L
-----				Majority
	730	740	750	760
148	-----			CARD4-Y CLONE
153	-----			CARD4-Z CLONE
701	RELQPCPSRLTYLRLSYNQITDGGVHVLSEELTKYKITY			CARD4L
-----ECX-----				Majority
	770	780	790	800
148	-----			CARD4-Y CLONE
153	-----			CARD4-Z CLONE
761	LSLYNNQITDVGARYTTKILD	ECR		CARD4L
-----WXXXXXXXXXX-----				Majority
	810	820	830	840
148	-----			CARD4-Y CLONE
155	-----			CARD4-Z CLONE
551	ARLGKKNKITSEGGKYLALAVHNSKRSISEVGH	GNQVGDEG		CARD4L
XX				Majority
	850	860	870	880
149	-----			CARD4-Y CLONE
155	-----			CARD4-Z CLONE
541	AKAFAEA	PNHPSLTTL	SLASNGISTEGGKSLARALQQNT	CARD4L
XX				Majority
	890	900	910	920
149	-----			CARD4-Y CLONE
151	-----			CARD4-Z CLONE
531	ELSLI	WL	TQNELNDEVAESLAEMLVNQTLKHLWLIQNOI	CARD4L
XX				Majority
	930	940	950	960
149	-----			CARD4-Y CLONE
154	-----			CARD4-Z CLONE
501	TAKETAGLADALQSWTETETICLNGNLKPEERKVVYEDER			CARD4L

FIGURE 14 (3 of 4)

20/58

<u>XXXXF</u>	Majority
149	CARD4-Y CLONE
154	CARD4-Z CLONE
161	CARD4L

Decomposition (Decomposition #1): Shade (with solid black) residues that match the Consensus exactly.

21/58

TCACGCGTCCGCGGACCCGAGCGGTAGCGCCCTCCCTCCCAGCTGTTGTCCCGCCCGAT  
CCGCGACCCCTAGTCCCGCGATCCCTTGCTGAGAGTCACCGTACTCCAGGGCCAACTGAG  
CCAAAGTCCTGCCAACTTGGGTGAGCAATGAAAGGCAGGATCCTGGGTGGTGGCCCTGAA  
TCCTGATTTGTCTGCCCTGCCAGCGAGACACATGTGGTCAAAGATGAATTTGAGAAAAGT  
AGCTGCTGGCTACTTGAACAATGGAGGAACACGGCCATCATGAGATGGAAGGCACCCCAT  
TGGGTTGTCACTCCACATTAAACTGCTGAAGATCAACAGGGAACATCTGGTCACCAACA  
FTCGGAACACTCAGTGTCTGGTGGACAACCTTGCTGGAGAATGGCTACTTCTCAGCCGAAG  
ATGCAGAGATTGTGTGTGCCTGTCCCACCAAGCCTGACAAGGTCCGAAAGATCCTTGACC  
TGGTGCAGAGCAAAGGCGAGGAGGTGTCTGAGTTCTTCCTCTACGTGCTGCAGCAGCTGG  
AGGATGCTTACGTGGACCTCAGGCTGTGGCTCTCAGAAATTGGCTTCTCCCTTCCCAGC  
TCATTCCGACCAAACTATCGTCAATACTGACCCAGTAAGCAGGTATACCCAACAGCTGC  
BACACCAACTGGGCCGCGACTCCAAGTTCATGCTGTGCTACGCCCAGAAGGAGGACCTGC  
TGCTGGAGGAGACCTATATGGACACACTCATGGGGCTGGTAGGCTTCAACAATGAAAACC  
TGGGCAGCCTAGGAGGCCTGGATTGCCTGCTGGACCACAGTACGGGCGTCTCAACGAGC  
ATGGCGAGACTGTCTTCGTGTTCCGGGACGCGGGAGTGGGCAAGTCCATGCTGCTGCAGA  
GGTTGCAGAGCCTCTGGGCGTCAGGCAGGTTGACCTCCACAGCCAAATTCTTCTTCCACT  
TCCGCTGCCGCATGTTTCAAGTCTCTTCAAGGAGAGCGACATGCTGAGTCTGCAGGACCTGC  
TCTTCAAGCATTTCTGCTACCCGGAGCAGGACCCCGAGGAGGTGTTCTCCTTCTTGCTGC  
GCTTTCCCCACACAGCGCTCTTCACTTTTACGGGCTGGATGAGCTGCACTCAGACTTCG  
ACCTGAGCCCGGTGCCGGATAGCTGCTGCCCCCTGGGAGCCGGCTCACCCTCTGGTCCTGC  
TGGCTAACCTCCTAAGTGGGAGGCTGCTCAAGGGTGCCGGCAAATTGCTCACTGCTCGCA  
CAGGCGTGGAGGTCCCCCGCCAGCTCTGCGCAAAAAGGTGCTGCTCCGGGGCTTCTCCC  
CAAGTCACCTGCGCGCTATGCCCGCCGGATGTTCCCCGAGCGCACAGCGCAGGAGCATC  
TGCTGCAGCAGCTGGATGCCAACCCCAACCTCTGCAGCCTGTGCGGGGTGCCGCTCTTCT  
GTTGGATCATCTTCCGTTGTTTCCAGCACTTCCAGACGGTCTTCGAGGGCTCCTCTTCAC  
AGTTGCCGGACTGTGCTGTGACCCTGACCGATGTCTTCTGCTGGTCACTGAGGTGCATC  
TGAACAGGCCCGCAGCCCAGCAGCCTGGTGCAGCGCAACACGCGCAGCCCGGCGGAAACCC  
TACGTGCAGGCTGGCGCACGCTGCATGCGCTGGGAGAGGTGGCTCACCGAGGCACCGACA  
AGAGCCTCTTTGTGTTTGGCCAGGAGGAGGTGCAGGCGTCAAGCTGCAGGAAGGAGATC  
TGCAGCTGGGCTTCTGCGGGCTTTGCCCGATGTGGGCCCTGAGCAGGGCCAGTCTTACG  
AATTTTTCACCTTACGCTCCAGGCCTTCTTACCGCCTTCTTCTTGGTAGCAGATGACA  
AAGTGAGCACCCGGGAGTTGCTGAGGTTCTTTCGAGAATGGACGTCTCCTGGAGAGGCAA

FIG. 15 (Page 1 of 3)



22 / 58

CAAGCTCGTCCTGCCATTCTTCCTTCTTCTCCTTCCAGTGCCTGGGCGGCAGAAGCCGGT  
TGGGCCCTGATCCTTTTCAGGAACAAAGATCACTTCCAGTTCACCAACCTCTTCGTGTGCG  
GGCTACTGGCCAAAGCCCCACAGAACTCCTTCCGGCAGCTGGTGGCCAAGGCTATCCTGA  
GGAGGAAGCGCAAGGCCCTGTGGGCTCACCTGTTTGCTAGCCTGCGCTCCTACTTGAAGA  
GCCTACCTCGGGTCCAGTCTGGAGGCTTTAACCAGGTGCATGCCATGCCACATTCTCTGT  
GGATGCTGCGCTGCATCTATGAGACGCAGAGCCAGAAGGTGGGGCGCCTCGCCGCCAGGG  
GCATCAGTGCGGACTACCTCAAGCTGGCCTTTTGCAACGCTTGCTCTGCGGACTGCAGCG  
CCCTGTCTTCTGCTCTGCATCACTTCCACAGGCAGCTGGCCCTAGACCTGGACAACAACA  
ACCTCAATGACTATGGCGTGAGGAGCTGCAGCCTTGCTTTAGCCGTCTCACGGTTATCA  
GACTCAGCGTCAACCAGATCACCGACACGGGGGTGAAGGTGCTATGTGAGGAACTGACCA  
AGTATAAGATCGTGACGTTTCTGGGTTTATACAACAACCAGATAACTGATATCGGAGCCA  
GGTATGTGGCCCAAATCCTTGATGAATGCAGAGGCCTCAAGCACCTTAAACTAGGGAAAA  
ACAGAATAACAAGTGAGGGCGGGAAGTGTGTGGCTTTGGCTGTGAAGAACAGCACCTCCA  
TCGTTGATGTTGGGATGTGGGGTAATCAGATTGGAGACGAAGGGGCAAAGGCCTTCGCAG  
AGGCATTGAAGGACCACCCCAGCCTGACCACTCTCAGTCTTGCAATTCAATGGCATCTCTC  
CGGAGGGAGGGAAGAGCCTTGCGCAGGCCCTGAAGCAGAACACCACACTGACAGTAATCT  
GGCTGACCAAAAATGAACTTAATGATGAGTCTGCAGAGTGCTTCGCTGAGATGCTGAGAG  
TGAACCAGACGCTACGGCATTATGAGCTGATCCAGAATCGCATCACAGCCAAGGGGACAG  
CGCAGCTGGCGAGGGCACTGCAGAAGAACACAGCCATAACAGAGATTTGTCTCAATGGAA  
ACTTGATTAAAGCCCCAGGAGGCCAAAGTCTTCGAGAATGAGAAGAGAATCATCTGCTTCT  
GACGGACGCTCCTGGGCAGGATCTTTGTCTAGGTTGCTCCTCAGTCACAGACAGCACTG  
TGCAGTCAGCAGGGTAGCAGGATGCTGTGCAGCGCCTGCAGCAAGGTGCCTGTCAGGAGC  
CCACACCTCCACAGTGACACCGATGTCCCTGCTCATGCTTGGAAGTGGTAGCACCCGCG  
CCGCGGCTGAGACCCTGCAGACGCAGGGAGTCTTAGGAACCATCGTCACCACTCAAAGCC  
AGCAGGGCATCTTCTGTACAAAGATCTCCCTGCATATCCACTAGACGGAAGCTGAAGGAA  
CGCAACAGCAGAGGAGGCCAACAGACGCCTGGCTGAAGGCTCCGTGGGACCAACGGTGTC  
ACCTTCAGAAAAGAGCTGGGAACTTGAGCAGAGCCGATGGTAACTTCTTGGGGAAAGAAG  
GCACCCAGTGACTGCATGGTTATTCTGAGTCCTCCTTCTCTGCTTAGTCCCTCTCACTG  
TACAGGTCTGTTTCTTCTCGCAGCTGTGGCTGCTGAAGTAGGTCCACTGTGGGGAGAGC  
TCATCACAGACTTTGGTTCGGTCTGGATTCTCAGTGGTGGCAACCGAGAGTCAGACGAT  
ACCTCTAGGTGAGTCTCAGAGGATCTCTATGCTGTGAGAGGGTTGAGGGCCACCCAGA  
ATTTTTTTTTTTTACCAGTTTTTACTGTGCCTGCCCCAGGAGGGAGAATTACTTCCCAGC

FIG. 15 (Page 2 of 3)

23/58

CTCCACAGCAGCAGGCATGGCTTGCCTCAATGGTCCTGAGATCCCAACAAACTCTCTCC  
CTTGCCTGTGAGCAGAAAGTATCTTCATGTCCTCAGAAGTTGGAGGGTGACTGGACACAG  
TTAAGACTCAGAGAGCCAGCTGATAGCTCAAAGCAAAGCATGGCACATACCCACCACCAT  
ACCATGGTGCGCATGGGATGGGACAGTTGGAATGTTGCAGATAACGTGTTCTTTTGCCAG  
TTCATTTGTTAATAAAATATTTAAAACGTTAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCG  
G (SEQ ID NO:42)

FIG. 15 (Page 3 of 3)

24 / 58

MEEHGHHEMEGTPLGCHSHIKLLKINREHLVTNIRNTQCCLVDNLLNGYFSAEDAEIVCA  
CPTKPKVRKIILDLVQSKGEEVSEFFLYVLQQLDAYVDLRLWLSEIGFSPSQLIRTKTI  
VNTDPVSRYTQQLRHQLGRDSKFMLCYAQKEDLLLEETYMDTLMGLVGFNNENLGS LGGL  
DCLLDHSTGVLNEHGETVVFVFGDAGVGKSM LLQRLQSLWASGRLTSTAKFFFHFRCRMFS  
CFKESDMLS LQDLLFKHFCYPEQDP EEFVSFLLRFPHTALFTFDGLDELHSDFDLSRVPD  
SCCPWEPAHPLVLLANLLSGRLLKGAGKLLTARTGVEVPRQLLRKKVLLRGFSPSHLRAY  
ARRMFPERTAQEHL LQQLDANPNLCSLCGVPLFCWIIIFRCFQHFQTVFEGSSSQLPDCAV  
TLTDVFLLVTEVHLNRPQPSSLVQRNTRSPAETLRAGWRTLHALGEVAHRGTDKSLFVFG  
QEEVQASKLQEGDLQLGFLRALPDVGPEQGQSYEFFH LTLQAFFTAFFLVADDKVSTREL  
LRFFREWTSPGEATSSSCHSSFFSFQCLGGRSRLGPD PFRNKDHFQFTNLFVCGLLAKAR  
QKLLRQLVPKAILRRKRKALWAHLFASLSYLKSLPRVQSGGFNQVHAMPTFLWMLRCIY  
ETQSQKVGRLAARGISADYLKLAFCNACSADCSALS FVLHHFHRQLALDLDNNNLNDYGV  
QELQPCFSRLTVIRLSVNQITDTGVKVLCEELTKYKIVTFLGLYNNQITDIGARYVAQIL  
DECRGLKHLKLGNRITSEGGKCVALAVKNSTSIVDVG MWGNQIGDEGAKAF AEALKDHP  
SLTTLSLAFNGISPEGGKSLAQALKQNTTTLTVIWLTKNELNDESAECFAEMLRVNQTLRH  
LWLIQN RITAKGTAQLARALQKNTAITEICLNGNLIKPEEAKVFENEKRIICF  
(SEQ ID NO:43)

FIG. 16

25/58

10 20 30 40 50 60 70 80 90  
 100 110 120 130 140 150 160 170 180  
 190 200 210 220 230 240 250 260 270  
 280 290 300 310 320 330 340 350 360  
 370 380 390 400 410 420 430 440 450  
 460 470 480 490 500 510 520 530 540  
 550 560 570 580 590 600 610 620 630  
 640 650 660 670 680 690 700 710 720  
 730 740 750 760 770 780 790 800 810  
 820 830 840 850 860 870 880 890 900  
 910 920 930 940 950 960  
 970 980 990

FIGURE 17

26/58

gatcatcggttactgcagccttgaactcttgtgctcatgtgatcctcctgccttagcctcccaa  
tagctgggactacaggtgcgccaccatgcctggctaattttttttatttttgtagagatgggtgt  
ctcactatggttgacaggttgggtctcaaactactggccttacttcaagctatctacccatctcag  
cctcccaaagcgctgggattacagtcattgagccaacttgctggccagataaaggctcttaagcat  
gggtccttctgtcttaggttagagaaacccacacacagctgggaggtggggtgagctcttctgt  
agcttttgcttctgtgatgatgtcattgatctctcaggggtgcgcagagtagcaggggcccctg  
gagggcgcgccctgaatcctgatggccttctgtgagaggacacacgcagctgaagatgaattt  
gggaaaagtagccgcttgcactttaactatggaagagcagggccacagtgagatggaaataatc  
ccatcagagctctacccccacattcaattactgaaaagcaatcggaacttctggtcactcacat  
ccgcaatactcagtgctggtggacaacttgctgaagaatgactacttctcgccgaagatgcgg  
agattgtgtgtgcctgccccacccagcctgacaagggtgccccggggacagggacgggcatggcat  
tgtgtggaccccgaggctagaagaggcctctccctgctgatctgagtgaaagcgtgggaggtt  
agtcacagcgggagggctgcattttggggtaactaatagcacacaaatgcctgggttagcaggttg  
cacagtcaggtattttacttctgtgtttgtgtctggagcaaacctgacatctcagttctcattg  
ctgtgtgtatttggttcccagacacttcattttagatcccctttaaattaggagggaaaaagaa  
ataagcataagagcatccccagcagcagtggtcattcagtgccctgaaggctggagggtgctt  
gttgcctgggtgagactcggagggggaaccgactcaggggtcaggaatgatgacatcccaggtgggt  
ccacagtgagaatcttccccgctccactgtgggacgccttaacagcccttacttccacttacgc  
tttgcggttatctcctgaaaaataaaatggagaccacaaatctcttctgggttagaggaatgacac  
aactcatttatgacatgaccccgctgggactcagaagagaccaggacgggttctgggggaagcag  
tagcacactcgtgtgctttgttctcttctcttgatttgtttccacatttttaacaagaaaaaa  
agccggttttaatatatggcctatcgccctcctactgtgtggccaggtgcctacctcatatgc  
caagggggtggttctcacctctccactctcattcctgcacagcaggttgtgtcaggttaagagggga  
caaggagaaggctgggcaccgtgggtcacgcctgtaatcccagcactttgggaggccgaggcagg  
cagatcacctaaggctcaggagtttgagaccagcctggccaacatggggaaaaccgctctctaata  
aaaacacaaaaattagtcgggcatgggtgggtgggtgcctgtaatcccagccacttgggagggtgag  
gaaagagaattccttgaacctgggaggtggaggttgacagtgagccaagatttgtgccattgcactc  
cagccctccagcctgggtgacagagcaagactctgtctcaaaaaagaaaaaaagagaggt  
agagaagtccatggctatttctgtctccttttatttttaggctcatggaagcctcctggttct  
tagagctgtggttcttatttctgtcaggaggtcatttccacagatttctgggctccaatatgt  
tgagctgcacagcagctggggggatggcatagctaccggctgtaactaagaactcagagccctgcc  
ctgagcctgcctgagggtccttatggtaggaggatgcccctcatgccagcccgctgcctcatgct  
tgtgtcacctccaggtccgcaaaatcttgacctgggtacagagcaagggcgaggaggtgtccgag  
ttcttctctacttgctccagcaactgcagatgcctacgtggacctcaggccttggtgtgctgga  
gatcggttctcccccttccctgctcactcagagcaaagtcgtggtcaacactgaccaggttagga  
gtcagcccccagcaagaccgcaggcaccagtgcaacagggccctggggggtttggttaattggctgg  
gccagccctgagtgccacctcaggaagcagggccaggtgctattttgattttagaaaggaacagc  
tgaatcctgtctcccaagtgcagcccaggtgggtgcgattgaaactgccacacctcgatgggtctg  
gtttatagaggggcttttgaagtatgggaatggcctgtgttctgaccccttgctttcttctat  
tctgacatatgtagacattttaatgggtgcacaaatcaagggtgtatttttttttcttttaaaa  
aaatcttttagctggacatggttagcacacacctgtagttccagctactcaggaggctgaggcaaga  
ggactgcttgagccccagagcttaaggctgcagcagagctatgatgtgtcccctacactccagcct  
gggtgacagagtgagaccctgtctctaaaaaaggaaagaaaaaaataaaaagccttgccaggtt  
tgattctaggcaagatttctgtcaccgttgagtgcagtccttatttccaaactaatggaagac  
cccatcagtttaactgatttagttcaataagattttttgtgtatccaccacatgccaaagacccta  
cactgtgctggatgtcaggagagacagtggtgagcagacacagacaggggttctgcctcaggggag  
cttcaagtcagctggaagagaccaccagtcagcaatctcaaaaatgtgtcaggacagcggcagtc  
caaggcatgtgagaacatatcattagggccaggatctgtctggggcaggagcttcttcttccctg  
cttttgaactctccactttgagacagctgttggttaacataccagcaccaaggacctaaagtcctgc  
cttttaagaatccaatatgttgttggaacagaaagcacaagacaggtgtgtgcttaggggaaac  
aaggccagccggcagagtgctcagtgctaggctccagcttccacagccctgcaggtgctgcccag  
ccactgctagcttctgactctgtctgctccttctgtctcccttggttcttcccccttagaaa  
aaaaagaaagatttccatagaggaattcttcttgcgaagacttctctgttggttccgttagcca  
gctactttactagcttttacagtgtaattcactctacaagcagctctcacacaaaagactacatat

FIG. 18 (1 of 10)

tgtatgattctgtttatatgaaatgtccagaaaaggtaaatctatagacaaagcaaatacagtagt  
tgcctacggcccagggtattggctacaaataggctccagaaaactctgggaagatggtagagatgt  
tctagacctggactgtgggtgagggttgcacaacttgtaaacttactaaaaattactgacaaata  
tataacactccctaacactttgggaggccgaggtgggcagatcgcttgaacccaggaatttgaga  
ccagcctgggcaacatggcgagaccccgctctctacaaaaaacacaaaaattagttgggcttggt  
ggcatatgcctgtgtcccagctacttgggaggctgaggtgggaggattgcttgagcctgggaggt  
tgagactgcatgattgggtcactgcaccctagcctgagtgacagagcaggaccctatctctaaca  
acaaaaaagcagtggttgggtggaggaggccagcgtggccatctggcctggccctcgagtgcgagg  
ggcttcagtggttagctgcagttcagtgatgacactgtgcggaggaataaggggtggcctgtctca  
gacactgatcccagctgaagtttgtcaccttcttctggcaaatctgaggtcaagcagagagatc  
aaagcctggggccctcagggtcaggaatgctggctctgtgacgctcccaggtcctgcatctgag  
gagtggtgcgctggcctcaggggcccaggttgtgaatttgtttatgcactcgccctctcctctt  
gagacctcctgtttgatgctgtttctgcctctctcctcaccctgctgctgtgcccctgccacccc  
ctcctccagtgagcaggtataccagcagctgcgacaccatctgggccgtgactccaagttcgt  
gctgtgctatgcccagaaggaggagctgctgctggaggagatctacatggacaccatcatggagc  
tgggttggcttcagcaatgagagcctgggcagcctgaacagcctggcctgcctcctggaccacacc  
accggcatcctcaatgagcaggggtgagaccatcttcatcctgggtgatgctgggggtgggcaagtc  
catgctgctacagcggctgcagagcctctgggcccagggccggctagacgcaggggtcaaattct  
tcttccactttcgctgccgcatgttcagctgcttcaaggaaagtgcagggctgtgtctgcaggac  
ctgctcttcaagcactactgctaccagagcgggaccccagaggaggtgtttgccttctgctgcg  
cttccccacgtggccctcttcaccttcgatggcctggacgagctgcactcggacttggacctga  
gcccgtgctgacagctcctgcccctgggagcctgcccaccccctggcttctgctggccaacctg  
ctcagtggaagctgctcaagggggctagcaagctgctcacagcccgcacaggcatcgaggtccc  
gcgccagttcctgcggaagaagggtgcttctccggggcttctccccagccacctgcgcgctatg  
ccaggaggatgttccccgagcgggcccctgcaggaccgctgctgagccagctggaggccaacccc  
aacctctgcagcctgtgctctgtgcccctcttctgctggatcatcttccgggtgcttccagcactt  
ccgtgctgcctttgaaggctcaccacagctgcccagctgcacgatgacctgacagatgtcttcc  
tcctggtcactgaggtccatctgaacaggatgcagcccagcagcctgggtgcagcgggaacacacgc  
agccctgctggagacctccagcgggcccggacactctgtgctcgctggggcaggtggcccaccg  
gggcatggagaagagcctcttcttcttccaccaggaggaggtgcaggcctccgggctcgaggaga  
gagacatgcagctgggcttcttgcgggcttctgcggagctgggccccgggggtgaccagcagtc  
tatgagtttttccacctcaccctccaggccttctttacagccttcttctcctgctggacgacag  
ggtgggactcaggagctgctcaggttcttccaggagtggtgccccctgcgggggcagcgacca  
cgtcctgctatcctcccttcttcccggttccagtgcttgcagggcagtggtccggcgcggaagac  
ctctcaagaacaaggatcacttccagttcaccaacctcttctgtgcgggctgttgtccaaagc  
caaacagaaactccttcggcatctggtgcccgcggcagccctgaggagaaagcgcaaggccctgt  
gggcacacctgtttccagcctgcggggcttacctgaagagcctgccccgcgttcaggtcgaaagc  
ttcaaccaggtgcaggccatgcccaggttcatctggatgctgcgctgcatctacgagacacagag  
ccagaagggtggggcagctggcgggccaggggcatctgcgccaactacctcaagctgacctactgca  
acgctgctcgggcagctgcagcgcctctccttctgctcctgcataccttccccaaagcggtggcc  
ctagacctagacaacaacaatctcaacgactacggcgtgcgggagctgcagccctgcttcagccg  
cctcactgttctcaggtgaggctgccaggcaaggggagcaacaggtgggcccgggcccaggct  
cggagggcacoggggaatggcatcatggaccaggatccccaggactcatgacctggcccttgga  
atgtccagaccttttcttcttagcagggcagaggtcaaggtgcaaagcttcgaggcaggtggac  
ctggatcagccacagctgggtgcccctgaacaaagtgttcaactctcagagcctccacgcctca  
tctggaaaaagaagatgctcataatcctatcaattatggccacagggaaccaatgttagttgagaa  
tgggtgaagtgcattacaaatattacctaattggaatgctcttacaacctgtaacttaggtact  
gttattgtctctattttggcagataaggaagtagaggcacagagaagttaatagcttgctttagg  
tcacacagctcagacatagcagtgccagaatgcataaagaaccttccctttaagattaatgtaag  
gctccgagatagccctcaaaaagtttctggaatatgggagcttttattactgcagagaaagcaga  
ccttggtgccagttggcactgggtgacttctgtgatcaacgctagcagcccttcacactgctagag  
acctcagttaaaatgctgactcggtgtgtttctgttccatagtttacgggaaacagagccca  
gtctgttttcttctattagcatttccatgttaaaataaaccttgtaaatctctacagggggttaa  
atttgccattacttgactcatgcatcttcaaaaagcagtagggatttggaaactgactcccagtg

FIG. 18 (2 of 10)

28/58

ctgtcacaccagtgtcagagtgtaaataattgcatggggacatgggggtgcagggggtcgaaggct  
gccctagcctgggaattggaaaacctggagtctgttctctgtactctcagccagtgtactctccct  
ctgtagccccaggcagtcctcacactcagtgccacctctgtccatctttttttttctccccaa  
atggagtcccgtctgttgcccaggctggagtgagtggtgctcagctcactgcaacctcc  
gcctcctgggttcaagcgattctcttgccccagcctcctgagtagctgggattacaggcacacgc  
caccatgtccggctaaagttttttgatttttagtaggacgggggtttcccatgttgggcaggctg  
gtcttgaaatcctgacctcaggtgatccggcctcggccttccaaaatgttggggttacaggc  
atgagccgcccgcacccgacctctgtccatctttcaatgggaaactccacaccagtgtggtgg  
ccctgcccttctgtgtccccagggtgaagcttcttccacaccagtgcagaaaaaacagcttg  
taggaaagcagaggatatgggttaaccacgggaagcacactcagttctctggctgcatcagttagg  
attagtttttagctgagagcgaaaacccccaaatgttggtgagttacaagcttatttctctcatgta  
aaagtctagaggtaggttagttcaggactggtatggagtctccatgacctccggagcccaggctc  
tcttctgccttctgttctgccatcctcactaccggcttcccatcttgggccaagagggtgc  
tcaaactccagccatctagtgcacactctagctatcagtaagaaggagggaagattgagagc  
atgcctcaatcttttaagaacacttcttggtattactaattatattgctgcttagatttcagaa  
cttaatggtatgggcagaaattaatgagatggggcccagctaaaagatgggggaatctattgctaa  
gaaagtatagatatgggaatgtctagcagcctgtgctgtcttggtgctggccatgccatgtacat  
acacactatttcccagcaccaagctggggactctgagggaaagggtccagagtgtctgacttgat  
cattttgatgtggcctaaaaatcaagcttttaattgttcagccttttacttggtatcaaggctcag  
cttggtgggtctaatggggcccaaggcttggtgttctaaagtaagttttattggaacgcagccata  
cccatttatttacttactggctgcttcacactacacagttgagtagctgtgacagagaccacatg  
gcccacagagcctaaaaatatttgctgtctgacactttacagaatgacatgagcagctcctttga  
cagtggggactcacagccttttcagtgacaaaatcaggggttagcccatgtgtttctggatggggg  
aagctgttgccattttgggtataacagttcttggtgagacctgtccagcattttgcaggacaccta  
acatcattggccctgcctgcaagatgacagggcactccctcctccagtcacaaccactaaaagca  
gcccctgacatttccaaacccatgccctccaccatacgagaaccagggtacaggggtctggctgaca  
cataggtcacacgcaaagggtggatgtcagaggtggctggcctcacacgtcctccctgtgtcctt  
cacggctcgtgtgaggagccaggggctgtgctgcagcctcgctcatgggctgggtgcaggatgggtc  
tgggcgccccacgttgggcaggctttgtaaggggctattggctgattgctgtggccatttccca  
ggggcgctctatacctgaaactccaggcctcagggcttctggatccttgtaagatttaagggt  
ccttcataatgagtgccctgactcgtaattttttgctgttttatttcagactcagcgtaa  
accagatcactgacgggtggggtaaagggtgctaagcgaagagctgaccaaatacaaaattgtgacc  
tatttggggatgtcttctccagaacactgggccaactacctagtaataatacagagctgcagg  
gaattcacattcccataggtccctggatgatcggcacggatggcccaggggtgggaagagcgctg  
gcccaggagttgagagtctgggttctcttctgtggctcggccagtcataaggtcttgctgagcct  
cagcctcctcacctgtaaaaactgggatcccagtaggcaagtaggcttacaactggttattggg  
ggatgcaacgcataataagggatataattataaatgctagaatcctgtttacatattagctc  
ggactattttgggtccataaatccctcatccagagcctttggggcaagaccgaatggggattctg  
agtgcattgctatggcatgacgtggccgcaggggtctaaggcagtgccccattttcaaacacttct  
atatttctcccgagaatgtatgaaacagtcacaaacagtggtggaagaaagactataagtagct  
ccacatcagttgccccaaagaattgtgagaaactttgggcattcagagcctttgagggttttgagg  
ctgagagaagggtatgcgggcccagccccacacaactggtggctctgcaagctggagcagttgttc  
agtttcttggggcctcagtgcccttcgatgttaatgaggacatggacgcaaacgacccccgggcca  
cactcggctccagggtctgtgtggctgtggaaccttggaagcctgagcttagctgcctttcaac  
ttccatctgtgtactattgaattggcattgagcggtgagatggctgaaaggtagacatcgagaa  
gttttaatatccagaatcttttcttctcaagacgctgaatgtaatcttagttgtaaataccatc  
acctgccagtcaccgagcactcatgcaccagggttctgcgttatgtcctaagatcctcataacca  
ccctgcaaggggactatcatcattacctctgtattacagatggagaaactgaggcacagagaggt  
aacgtgacttgtctcaggccataaaagctggggaaagtagtgagctgggtttgaaacctgagctgt  
gagacctcagagccctaaactctgggtgcctgtgtgttcccccttcaacccagactttggaaatca  
gtagacaccatattgcttcaaaaaacaggggctattaaaatgacatcaggagccagaaagctcat  
ggctgtgcttctcttgaaagttatacaacaaccagatcaccgatgtcggagccaggtacgtcac  
caaaatcctggatgcaatgcaaggcctcacgcatcttaagtaagtggggtaggcaccaggtcct  
tagtatattctctgatcaccccttctgtgttcaagattaaatgtcacagtaagagcttct

FIG. 18 (3 of 10)

29 / 58

atcctaaagccttccacttgtcccagggccatgttgggtcaagtaagatacctctgtgtgatctg  
tgaggcttggattctggaagggcctcccgttattggtaggggaaagggttggcattttgatttca  
ttaactactaggccgaagaaaggactaactctcaccctttctgggtggtctttttgccccaaagga  
gtttcctgtcgggttgcaaggaagagcttgggcccttgccctgctgtaggtgtgccctgcgcagg  
gggtgacagtgcgccaggcttggagcctctggtcctgccctgacagtggccacataccttgaccc  
ttggcagtc aaagtgggacctcccaggtctcccaggggaagtacgtgatgctgctgaggtcaatt  
agaggaccccagggaggggtcaggtccctgagcttctgcagagactgtggaccatctcctggaga  
ggaaccctgactgactgtcctcagggcttcagttccctccctgacaggaggcccaggccatgggt  
cttgtggatcccagaagaaagtgtacggttcccaagatggggctggaaggggctctgtgctgggg  
aggaggtgacccacattggagccccctgcatagctggaggtgactgtgtgtgactctctctgca  
gactgggaaaaaacaaaaataacaagtgaaggagggaagtatctcgccctggctgtgaagaacgc  
aaatcaatctctgaggttgggtgagtagaaggggatggatgtatgtggtacaacctgctgtgtgt  
gtggggggcgggccttgctgttcttttcatacatcagtacaccagaaggaccactggggctcgct  
gtcggggagagatagtggagagctttcaccatgctgcgaaactgaaaccgtgcccat taagcaat  
aactccccgggtccccctccccctgctcttgcagccacctgctacttactctctctatggttt  
tgactactctacctcatgtaagtggaaatcatacagtat ttgccttttggggatgggtgatttcac  
tagcatctgtcctcaagattcgtccacatgggaagcatgggacaggatttcctttttttttttt  
tttttttttttttgacagagtctcgctctgtgcgccaggtggagtgacgtggcatgatctcgg  
ctcactgcaacctctgccttctgggttcaagcgattctctcgccctcagccacacgagtagctggg  
attataggcaccgcgccaccaatcccagctaatttttgtatttttagtagaggcggggtttcacca  
tggtggccaggctggtctcaaaactcctgacctcaaatgatccaccacctcggtctcccaaagt  
tcaggattataggcgtagccacctgccccgccaggatttccttctttttaaggctgagtaat  
actccattgcatggctatgccacattttgtttactcattcatccaagaacagacactggcttgct  
tctatgctttgggtgtgtgaataatgctgctgtgcacatgggcatacaaatgtctcttcaagga  
ctgccttcaattcttt  
agtttttagggtacatgtgcacattgtgcaggttagttacatatgtatacatgtgccatgctgggtg  
cgctgcacccactaatgtgtcatctagcattaggtatatctcccaatgctatccctccccctcc  
cccgaccccccacacagtcctccagagtgtgatattccccttccctgtgtccatgtgatctcattgtt  
caattcccacctatgagtggagaatatgcggtgtttgggtttttgttcttgcatagtttactgag  
aatgatgggtttccaatttcatccatgtccctacaaaggatatgaactcatcttttttatggctg  
catagatttccatgggtgtatgtgccacattttcttaatccagtcctatcattgttggacatttg  
gggtgggttccaaagtctttgtctattgtgaatagtgccacaataaacatacgtgtgcacgtgtgt  
atagcagcatgatttatactcatttgggtatataccagtaatgggatggctgggtcaaatggta  
tttctagttctagatccctgaggaatcgccacactgactccacaatgggtgaactagtttacag  
tcccaccaacagtgtaaaagtgttccatatttctcgcacacctctccagcacctgctgtttcctga  
ctttttaatgattgccatttcaactgggtgtgagatgatctcatagtggttttgatttgcat  
ctctgatggccagtgatgatgagcatttcttcatgtgttttttgggtgcataaatgtcttcttt  
gagaagtgtctgttcatgtccttcgcccacttttgatgggggttgtttgttttttcttgtaaat  
ttgtttgagttcattgtagattctggatattagccctttgtcagatgagtaggttgogaaaattt  
tctcccatggtttaggttgccctgttcaactctgatggtagtttcttttgctgtgcagaagctctt  
agtttaattagatoccatttgtcaattttgtcttttgttgccattgcttttgggtgttttgacat  
gaagtccttgcccacgcctatgtcctgaatggtaatgcctagggttttcttctagggtttttatgg  
tttttaggtttaacgtttaaatctttaatccatcttgaaattgatttttgtataagggtgtaaggaag  
ggatccagtttcagctttctacatatggctagccagttttccagcaccatttattaaataggga  
atcctttcccattgcttgtttttctcaggtttgtcaaagatcagatagttgtagatatgcggca  
ttatttctgagggctctgttctgttccattgatctatatctctgttttgggtaccagtaccatgct  
gttttgggttactgtagccttgtagtatagtttgaagtcaggtagtgatgcctccagctttgtt  
cttttgggttaggattgacttggcgatgcgggctcttttttgggtccatatgaactttaagtag  
ttttttccaattctgtgaagaaagtcattggtagcttgatggggatggcatgaaatctgtaaat  
accttgggcagtatggccattttcacgatattgattcttccctacccatgagcatggaatgttctt  
ccatttgtttgtgtcctcttttatttcccttgagcagtggtttgtagttctccttgaagaggtcct  
tcacatcccttgtaagttggattccttaggtattttattctctttgaagcaattgtgaatgggagt  
tcacccatgatttggctctctgtttgtctgttgggtgtataagaatgcttgtgattttgtac  
attgattttgtatcctgagactttgctgaagtgccttatcagcttaaggagattttgggctgaga

FIG. 18 (4 of 10)



30/58

cgatgggggttttctagataaacaatcatgtcgtctgcaaacagggacaatttgacttcctctttt  
cctaattgaataccctttatttccttctcctgctgattgccctggccagaacttccaacactat  
gttgaataggagcggtgagagaggggcatccctgtcttgtgccagttttcaaaggggaatgcttcca  
gtttttgccattcagtatgatattggctgtgggtttgtcatagatagctcttatttttgaaa  
tacgtcccatcaatacctaattttattgagagtttttagcatgaagggtgttggaattttgtcaaa  
ggctttttctgcatctattgagataatcatgtgggtttttgtctttggctctgtttatatgtctgga  
ttacattttattgattttgcgtatattgaaccagccttgcacccagggatgaagcccacttgatca  
tggtggataagctttttgatgtgctgctggattcggtttgccagtttttattgaggatttttgc  
atcaatgttcatcaaggatattggtctaaaattctctttttgggtgtgtctctgcccggctttg  
gtatcagaatgatgctggcctcataaaatgagttaggaggattccctctttttctatttgattgg  
aatagtttcagaaggaatgggtaccagttccctccttgcacctctggtagaattcggtgtggaatcc  
atctgggtcctggactctttttgggtggtaaaactattgattattgccacaatttcagagcctgtta  
ttggcttattcagagattcaacttcttccctgggttagtcttgggagagtgatgtgtcgaggaat  
gtatccattttcttctagattttctagtttattttgcgtagaggtgtttgtagtattctctgatgg  
agtttgtattttctgtgggagtcgggtggatgatatccctttatcattttttattgtgtctatttgat  
tcttctctctttttttctttatttagtcttgcctagcggctctatcaattttgttgatcctttcgaaa  
aaccagctcctggattcattgattttttgaagggtttttgtgtctctatttcttccagttctgc  
tctgatttttagttattttcttgccttctgctagcttttgaatgtgtttgtctcttgcctttctagtt  
cttttaattgtgatgttaggggtgtcaattttggatcttctcctgctttctccttgtaggcatttagt  
gctataaatttccctctacacactgcttgaatgcgtcccagagattctggtagtggtgtcttt  
gttctcgttgggtttcaaaggaacatctttatttctgccttcatctcggtatgtaccagtagtcat  
tcaggagcaggttgttcagtttccatgtagttgagcggctttgagtgagattcttaactcctgagt  
tctagtttgattgcactgtggctctgagagatagttgtttataatttctgttcttttacatttgct  
gaggagagctttacttccaactatgtgggtcaattttggaatagggtgtgggtgtgggtgtgaaaaaa  
atgtatattctgttgattttgggggtggagagttctgtagatgtctattaggtctgcttgggtgcaga  
gctgagttcaattcctgggtatccttgttgactttctgtctcattgatctgtctaatgttgacag  
tgggggtgttaaagtcctccattattaatgtgtgggagtcctaagtccttctgtaggtcactgagga  
cttgcctttatgaatctgggtgctcctgtattgggtgcataaaatatttaggatagttagctcctct  
tgttgaattgatccctttaccattatgtaatggccttcttctgtctcttttgatctttgttggttt  
aaagtctgttttatcagagactaggattgcaaccctgcctttttttgttttccattggcttgggt  
agatcttccctccatcctttttattttgagcctatgtgtgtctctgcacgtgagatgggtttcctga  
atacagcacactgatgggtcttgactctttatccaacttgccagctctgtgtcttttaattgcaga  
athtagtccatttatattttaaggttaatatgttatgtgtgaatttgatcctgtcatttgatgt  
tagctggcgattttgtctcatttagttgatgcatgttcttccctagtcctcgatgggtctttacatttg  
gcattgattttgacgggtgggtaccggttgttcccttccatgtttaccgcttccctcaggagctc  
ttttagggcaggectgggtgggtgacaaaatctctcagcatttgcttgtctataaagtattttattt  
ctccttcaactatgaagcttagtttggtgggtatgaaattctgggttgaaaattcttttcttta  
agaatgttgaatattggccccctctcttctggcttgtagggtttctgcccagagagatccgctgt  
tagtctgatgggttttccctttgagggtaaccggaactttctctctggctgccttaacatttttt  
ccttcatcttcaactttgggtgaatctgacaattatgtgtcttgaggtgtctctctcgaggagat  
ctttgtggcgttctctgtatttctgaatctgaacgttggcctgccttgctagattggggaagtt  
ctcctggataaatatcctgcagagtggttttccaaactgggttccattctccacatcactttcaggta  
caccaatcagacgtagatttggctctttcacatagtcocatatttcttggaggctttgctcattt  
ctttttattcttttttctctaaacttcccttctcgttctatttcttcttcttcttccattgc  
tgataccctttcttccagttgatcgcatcggtcctgaggcttctgcattcttccagtagttctc  
gagccttgggttttcagctccatcagctcctttaagcacttctctgtattgggtattcttagttata  
cattcttcttaaatttttttcaaagttttcaacttctttgctttgggtttgaatgtcctcccgtag  
ctcagagtaatttgatcgtctgaagccttcttctctcagctcgtcaaaatcattctccatccagc  
tttgttctgttgcgtgggtgaggaactgcgttcccttggaggaggagaggcgtctgcgttttagag  
tttccagttttctgttctgttttttccccatctttgtgggttttatctacttttgggtctttgatg  
atgggtgatgtacagatgggttttcagtgtagatgtccttctgggttggttagttttccttctaaca  
gacaggaccctcagctgcaggtctgttggaaatccctgcctgtgaggtgtcagtggtgcctctgc  
tgggggggtgcctcccagttaggctgctcgggggtcaggggtcagggaccacttgaggaggcaggt  
ctgcccgttctcagatctccagctgcgtgctgggagaaccactgctctcttcaaagctgtcagac

FIG. 18 (5 of 10)

agggacacttaagtctgcagaggttactgctgtctttttgtttgtctgtgcctgccccagagg  
tggagcctacagaggcaggcagcctccttgagctgtggtgggctccaccaggttcgagctccc  
ggctgctttgtttacctaagcaagcctgggctatggcgggccccctccccagcctcgttgccg  
ccttgacagtttgatctcagactgctgtgctagcaatcagcgagattccgtgggcgtaggaccctc  
tgagccagggtgtgggatatagctcctggtgctgctccttcttaagccggtctgaaaagcgcaata  
ttcgggtgggagtgaccgattttccaggtgctgctcctcaccctttctttgactcggaaagggga  
actccctgagctccttgccgttcccaggtgagcgaatgcctcgccctgcttcgggtcgcgacgggt  
gcgcgcacacactggcctgcgcccactgtctggcgctccctagtggatgaacccgggtacctcag  
atggaaatgcagaaatcaccctgcttctgctgctcagcgtgggagctgtagaccggagctgtt  
cctattcggccatcttggtcctcctcccaattcttttgggtatatatccagcagtgggattgct  
ggatcacatggtaatttttaattttttgaagaatcatcatactgttttccacggcagcagcacca  
ttttatgttcccaccaacagttcattctagtttctccacatccttgccaacacttgctattttct  
ctttttgacagtacccatcctaagtgtgaggtcctgtctcattgtgggttttgattcttgagg  
cttttaagctttttgtttcattataattttttatggattacaaaaggaacacaggttaatttta  
tttggaaactatgaaaaataataaaaattatctctcagaaaatgattcttggttaacatttaagc  
tcagttaagctctctcactttctctcccttctctctctttgtacaacttttaaaaaatatagtag  
gggtgagactatatgtatctatactatagtaggggtgagactatatgtatccttcccttttccact  
taatctcatgccttgagtagctttccactttattaaaaatgtgatgccattcaattgtatagtaa  
atacatatatgtaagcaaacactgaaaactcttattctgggttccagcaagccatacctggaat  
ggtgtaagcaggtagtttgcttgggtgtgaacgtgttgttgaggcagctgccattgtgtgtgagt  
gggcccacgaacttgttctgttgtgtgtgtagacagtggtgtgctgattcttaggaacagccaac  
gctttgtgtgagccacacacaggttcttaagtgttctgcttctgttaactcagtgaatcctcacaac  
tccatgacggaatgctctaattatccccattttatagatggggcaactgaggtccaagagactac  
ataatttcccgaagttcacacaggttagcagatggcagagccgggtcaggagtccaccatcttacc  
acgcagactgttttagccagagactctccggatctgctgtaggggacagaatacagctttatcgc  
cgacctgtccaccaagatggcctgagccacagagcttgggttgggtaacgtcctctttatgtgac  
aggaacgttgcctgatggggtttctgaaggtacttccctgctctttgtctcctggaagactgtgtct  
tcaggaatgtctctgaccctgcccagagttgaacggatgctgggaaccagcactgcacacggc  
cttccctccagactctgcgcactctgtgctccacagagacatgcaggtgctttctctcatga  
gctcaggtcctgggctgacagctctccgaagctcgtgggtgaggtcgggtcttaactgtgcccac  
ttgccgatggcctctgttcacaaggcttccctgctcttcgatcttgcatcccccttgaatttg  
aaatccagagcagccactcagagaccagtgtaggaattagtgtccaggccacagatccagggga  
ctgggcacaaacatctgcctgttgagtaggaactgagctgtggccattggcaaaaaaggagggt  
gagcatggctgtttcttggggagctaacattcactatcttgtctcctccctcaggatgtggggca  
atcaagtggggatgaaggagcaaaagccttcgcagaggtctgcggaaccacccagcttgacc  
accctgaggttaactgtgctgctcctcaggggccaactggctccctccagctgctctagg  
tttgcctggggaagggtgattcgtgctcctaatagaagaggaatttgcatgtgtgattttccttac  
tctgtcaaacctttctttgatgcataagaggccatctagtaaagcacattcttctctttttta  
actttaagttctgggatacatgtagaagatgtgcaggtttgttacataggcaaatgcagccatg  
gtgatttgcctgacatcaacctgtcatctaggttttaagccctgcagtcattaggtatttgc  
ctaattgcttgcctcccttgccccccacccccacaggccctgggtgtgtgtgttccccctccat  
gtgtccatgtgttctcattgttcaactcccacttacaggtgagaacatgcagtggtttgggttttt  
gttctgtgttagtttgctgagaatgatggtttccagcttcacccatgtgccagcaaggacatg  
atctcattttttttatggttgcatagtattccatagtgtgtatgtgccacattttctttatcca  
gtctatcactgatgggcatttgggttgggttccaagtctttgctattgtaaatagtgtacataa  
acatacatgtgcttgtgtctttatagcagaatgatttataatcctttgggtaaataccagtaaat  
gggattgctgggtcaaatgggtatttctgggtctagatccctgaggaatcaccttaagtgtttatt  
cagctcagtgaaattctgcatgtgtcccacaccagccaaccaccccccatcaagacagaggaca  
tttccagccctcagccatccctgcatgtcccttgctggttagagggagggtttcctaagtgacaga  
tgaaacttaataagatgctggccagcagattcctgccccttcttgtcctcaggatgatgtgga  
aaagagggactcttctctctataaatggggatgcacctaccagcccccgcttaggctgctggc  
caaatcttgggaccttggtatgtccacggctctgctgctgttcttctaccactgaaaaagagtc  
caagaagggtggggacagtagcagaagagactttgccaggtcttgagatgggggtaccttgatggg  
gccagcctttagaaggacagcttgccaggcctcgccagcctcctgcccattgtgcagaaacctgag

FIG. 18 (6 of 10)

gtgccgacccccagcccaactgttgtgtgagcaggctgtgctgatgacccatttcccgtccagcctg  
cccttgtgctctgtgtgtgggctctggggcagcagcgcctgggcaactactgctgcagctgaacac  
ttctgcatcctgccccgagtgagcctgggctggggccacagccaggcagaggcttcccagctgtt  
ctgatgttgaagctaagattgaatgtagatgtgtctttaaataattcaccccaagtggtgttccctt  
ctagtcttgcgtccaacggcatctccacagaaaggaggaaagagccttgcgaggggcccctgcagcag  
aacacgtctctagaaatactgtggtaatagctcgagtcatttcatgtttgtttgttttctgt  
gatagggtcttgccttgtcgtccaggcttgagtgcatgtgtgtgatctcagctcactgcagcctc  
cacctcccaggctcattcgaacctcccgccttggccttccgagtcctgagactataggcatgcac  
caccacacccagttaattttaaaatttttgttagagatgggggttttgcctatgttaccaggtgg  
tcttgaactcctgggctcaagcagttctcctgccctggcttctcaaagccctgggattgcaggtg  
tgagccactgcacctggcagagtgatcttggagggtttaggtcccaggaattatcccagggggc  
tgcacatggcctggaatcttaacagaaaagggtgtctcccaattggaaaggctctaggccttctcag  
ttaagttgataatttctccttagagaagagaatagccacttctacaagcataaacagggtacagga  
ggaggaagtgaggctccgggagcctggatctgaggccttggccttctaggccccaggagaactaga  
acgtggccatgcaagctatccagggtatccttggatcccttcagatgtgcttagcagaggccaac  
ttccacacacttggctcaaaattttctcccttctcctctctcatctgccttccccagggcagcct  
cctccttccccagggtcttcacatcagggttggccttctatgctccatccagctcatctgtcactt  
gtcacctgaagcccacagtcctcgtcctcctctgcactctagggcacttactaagtggatgtgg  
cctcctgagagtggtttttgttgggtgttcccttttttatggccacttaagtgtttattttgcttt  
atgtgtattacatctctgtatcataaattccatacagggtggctgggagcagtgactcacatctg  
taatcccagtaacttggaaaggctgaggtgggaggatcgcttgaggccaagagtgcgagactagcc  
tgggcaatatagcgagaccctctatctacaaaaaaaaaaaaacattccttacagggttaagtgggg  
agttgtattacaacctccctatcatctactcagagcccagtgctcatttgatcttgctaaatta  
gttactgagaataatgacaatatcctcctcatgagagagtttgacattaggcctgctgtccagt  
aagtgcatttttaatttcttccccctcaacaaatcatttaacattttgaaaagtagtttatgtttt  
ttggaaaaaatgtaagacactaaaggaggacatgaaagtacctcctaaagttcctgctaaaagga  
ggaagtgaagtagctcccttgggttttccaaaataaccttcccttctagccttttgttctat  
gtatgttcaaagatatgcaaaacagaaatagcattcaagcagtggtcttaaaaaatattgtaatcac  
atactttacatgtctcctttaggggttctccatcttgatgctgttgacattttgggtccaagtgat  
tctttattatggtagggctgtcctgtgcatcatagacgggttagccgcatctctgccctgtacct  
cccagtggtgaggatcaaaaatgcctccggacatggccagggtgccccatggagagtgaaatcaca  
tggatagtagtaattgtcaacacctagaagccctcaagtgctgactgcagtcctatgtgtattcta  
cactttttccctgtgttaactcactcagcctcacaacctctatacgatctctactgttaacgt  
tcaccagtgagaaaaactcagacccaaagaacttaagcctgttgcccagggtcacctgctgggtgg  
gtgatataaaagtagtccagagctgagtcggagtagatgtcaatgctgtgttcttctccctctca  
ttctacctcatttctccctacaagctgcacaacatctcgaatagatatcacaatatatttcatcag  
ttgtttctgatctaaatttgttcagattttacattaggataataaccacaatgcagtgctgcaatgt  
ataaagctttgtgtgtatatccttgcacactgtagggttaaatttctagaagcttgattgtcttaa  
aatgaagcacattaaaaatttggggcaggcacatccaaactgcccttcaagggaatttttttttta  
aatgttctttctgttctattcttcttctaatgattctttcgtccactggcacaagtgggtccta  
ccctgtttacaccaaggagcttgggtgctttatccagaccacttctgggttctaaggaccattgag  
agacttctgaacttccagtcacttaacttgggtccctcacaagtttaactgagagcaaaagtactg  
aacacatttttaattgtgcagtcagtgactgtttcagggtcttcaaaactaacttggataaacacactgt  
cagtggtgttcaaggggaccctgggactagaggagaactgagaagcaggcatggcccttgtttt  
ccgtggggcccccatcttccatgaaatctgagggtcagcaaaagggtggggaggagggtgggtcc  
tctacaggtagctgggctaagaaataggagcccagggtacaggatttgcatataaaatgagtccta  
ttgaccttctgtggggctgacaggctgggcttggagcctggctgttttctgggttctcagcaagt  
gatcatctgcatagctggagagccttgggctgagctcccgctcctgtgaactctaaaacaatgtc  
tgccaagttaggctctcttgagtaaaatacttcttttttctttaggctgacccaaaatgaactc  
aacgatgaagtggcagagaggttggcagaaatgtgaaagtcaaccagacgttaagcatttatg  
gtaactcagagagccttacaatttccagactgtgctactttcaaaaagtattttttgagataaaaat  
ttacatactgtaaaattcactctcttaagtatacaattcagagggttttttagtgcaaccatcacc  
acctaatctagaacattttcactcctcctccccactccaaaaagccctgggtatccattaagcag  
tcactccctgtcctcctccccagaccctggcaaccactaatccgcttctgtctctatggatttg

FIG. 18 (7 of 10)

cctactctgggcatttcatataaatggaatcaagcaatatgtgaccttttgtctctgtgttctag  
catgtttcattcccttttatggctaataatgataattcactctaaggaaattttgcagtttattaat  
cagttgatgggacatttgggttgtttctactttttgactattatgcgtaatgctactgtgaacac  
tcctgttcatgcttttgggtgaacatatgttttcatctcttttgggaatatacctgggaatagaa  
tttctgggtcatatggcaattctgttaactttttgaggagccaccaaactgttttctaaagtggat  
gtactattttacattctcgcagcaatgtatgtggattccaatttctccacatcctcaccaacac  
ttattattgtccatcttttaaaatctagtatactagtggatgtgaagtaaatattgtggttttga  
tttgcatttccctgatgacaacaatgttgaatgtctttttatgtgcctactgggagtcgtatag  
cttctttggagaaatgtctccatatcctttgcccatttttaaaattgggtttgtcttctaagtctg  
agttataggggttctctatatattctgggtgctagacctttactagatacaggttttgcaagtat  
tttctttctttctgtggagtttttctctttcttgatagtgccttttaaaggacaacagttttta  
atttttgttttttttgagatggagtcctgtcttctgtcaccagacaggagtgagtgccatgatc  
tcagctctctgcaacctccacctcctgggttcaagcgattcttctgcctcagcctcctgagtagt  
tgggattacaggcatcagccaccatgcctgtctcattttgtatttttaatatagagatggggtttca  
ccatttaggcccaggctggtcttgaactcctgacctcaggtgatccacctgcctcagcctcccaa  
agtgtgggattacaggcgaaaagccactgcacctggccaatagttttttaattttgatgaagtc  
aatttatctatttttttctttggttgcctgtgctttcagtgcttctatctaagaaatgatgtccta  
atccaagatcacaaagaactccacctaaagttttctgttaagcggtatagttgtttccctccacat  
ataggtctgcaatccattttgagtttaatttttgtatagtgtaaagtgaggggttaacctcattctc  
ttgcacgtggatatccagctgtcccggcagcaccacgtgttgaaacagattatcttttccctattga  
atggccttgacaccttctgcaaaaatcaattgaccataaatgtatgggtttatttctgaattctc  
tgttctggtccattgatttatatgtctctcctatgccaggaccattgctgtagctttgtgtagta  
cattttgaaatcaggaggtgtgagttctactttgttcttcttctcaagattgtttagaccattc  
tgggttctttgcatttcttatgaattcagactcacccttgcaatttctgcaaaaagactagactc  
tgctacatatgttttttcttttcttttagcctgcagaattatttgatcccatccctaagtg  
aggccagcctctccaggggagagcagagctaggacagggtcagaaagagagtccttggtgctttgt  
gcattcccaacctgcactggccctagtgaaggcagcccagtggtggatgtgcctggacactgc  
aggctttttaggggcattaggtgctctccttccctggcctcctgccacatcttggttgagggtgc  
cttccctgccttcaaaaaagcctaagtgggtgactagaaaacagcagagtgtaactgaatacagaa  
cttggtgcccacttccctggttctatttttgtcccttttgaaagggaaggtcattacctctgccat  
tgaaccaggggcccttcccttggtgggtatgggtgggagcaccagatcctggctgcagccag  
ccaccagtggctcctgtgtgcttgggcagtaacagtgacaagagctcccttccccctggacactgt  
gcctaataacctcctccttgaaatctcacacaccagtggtatggggggcactcttatagttattct  
cagtttacagatgacacaactgaggcacagacagatgcgtttatttcttcaagggttctgtagctg  
aacagtggggagggaggggttaagaggagctgcaccgctctgcaatactgcctctcacgagggga  
gtcctcttcattcatgacagcatagggccctcgtcttccctggtaagggttcccttcttgggtcag  
tgccaggatttctaagggtcatgtttagcaggagcctattctacaaacagccaggagcagggaaat  
gactctgtgatgaagcggagacactacagcctcttgatgcatttatttccctggttgggttagaag  
cgtagctgcccaggagcatttcaggagaggcctggcttccctagcgatagctgaaaactttgtt  
tcatttgaatcactgctaccagaacaatggggtgcatttctcagagtcctccattattaaagcttt  
tccactgagccccatgagaactattcatgagaactatttcatggcagcataactgtttctcctcc  
ctccctcttgcatgtttggtagcctcttaactttaaaacctgccttgcccttccctagctacctgg  
aaggagacgtcagacttccctgtcccattggtgtgtttcttacaatttgttgttcagattggtgggtc  
tcccaaataatataaaaaatataaatggagtctcactctgtcaccaggtggagtgagtgaggca  
cgatcttggtcactgcacacctccacctccaggttcaagcaatttctcctacctcagtcctcccgag  
tagctgggagtagcaggtgcacaccaccatgccagctaatttttttgatttttaatatagagacag  
gttggtccaggatggtatcgatctcctgagctcgtgatccaccacctcgccctcccaaagtgcgtg  
ggattacagggtgtgagccactgcacccggccccaaatattttgattatgcacctctgcagtga  
aatgcaaacacacacatcagttcatgtattacattatgttctactataaaaacaaacagaaaattt  
aaaaaatatcaagctatcctttactctagtggatcttacctggacacttttagccagatacaaag  
tcacatggactcagttcttccctgaccaacttgtctcttatccaaaacaccttgcaactccc  
ttacgaaggggtcaaatttgatccagttatgtgattttatacaagttatgttcttcttccagge  
ttatccagaatcagatcacagctaaggggactgccagctggcagatgcgttacagagcaacact  
ggcataacagagatttggtaagatcccagcgtttgtcacagtaataacaccagtgactgttact

FIG. 18 (8 of 10)

caccaccactgactgtgcaaggcacaacgcaggggtggtttctgtttattcctccagcaacctgc  
acagtaatggtattacctctgttttacagaggttagacagaggcccagaccagtgaataaggttg  
cccaaggtcactacgagagaagctagaattcagcccagaatgcctgattccatattctgtgctct  
cctgccctgggccccgcctcatctaccttcattgggtgggagtgagggaagtggccagtgaat  
gatttccctagtggagtaaatccccctgggactcagcaattgagagatgactgtgtggccagga  
gtttggagctcattctccccctttctgggttccgtaagacatttccaggctgacttgaactgac  
ctgtgctctttgtctacttctttttctgctttgagaacttcttatgctaataagaagaaaaaa  
gtttgctttactgtgacattgagcgccatgccacttcttcttgctcccataaggcacagacac  
tccccactcagcagctcccttaacaacttaattgcctgggtgacgtgggactgggtggatgctgg  
gagaggggcttattaactatgtcctcttcatgactggggagaatttcatagccaattaaaaa  
aaaacaaaaaacagctccttggccaacacaggtcctcatacagtgtttttaaactttgcttta  
gaacttgtttggaacttgtcataaaatcgatcagtttgggtgaattgcaaccaacaatataaaa  
agaaaacagaacagaacaaaatatcaggatgcaatgtgcatggatgaaagtatcatttcattca  
tcttagttcatgcttgcagtgtagtggtgtgtgtttgcataagtgttgggtcacaacataaaat  
gtaattcttatttagggttgttagacaaaagggttttttttaaaaaaacactgttggctaggcat  
gggtggtcatgcctgtaataaccagcacttggggaggccaagatgggcagatctcttgagcacag  
agtttgagaccagctgggcaacatgcgaacccccgtcactacaaaaattagcccagacatgggtgc  
tatgtgctgtagtcccagctactcaggaagctgatgtgggaggatggatgcatgggagatcaag  
gctgcagtgagccaggatcatgccactgcaatccagcctgggtgccagagaccctgtctcaaaaa  
acaaaaagaaaaaaagaaaaacaccatcatagagaatagagcccagatctaaacagacacctgt  
ggcctgtgtgctgcaagcccagcctgcccagcagcctgggaagcactggaggggcactggaact  
gtttgcatgggtgtttgcccctcaggccactcgtttctgtgattctttaagttttgaggacagca  
ggcagagggggagaggaaggagactgcccagactacagaacagtttgagagcacagttggcttcc  
acttttctctgtagctgggtcaggcgggtagtaaaagacctacagttgctttaattctgtcaagtt  
caaaatctgcattgcttccctcttgagggtcaccattcctacacaaggaaccatttttagtagggc  
caggagacttcagcttcaaggcctgcacttgtgtcagggtggagaggggaactggccaccaattc  
agagaggggcaggacaggcggcatgggtgctgggtcttgggagtgtcttcacttaggtccctggctt  
gttctgggagcctccagagcatgctcctctgtgtgtgacttcatgggactgggctctgagaaggc  
tgttgcttgttggccctgccagggactgccacaccaggccacagggttgtgggttgagctggcgg  
gggagccacgttccaggagcagctctgttgggacacaaacttacagagtaagccttctccttgg  
acttgttaactgtactgacacttatttctacctcattccttctgaaaataacttggaagtctga  
agtcccttgatgagttctgtctttaagaacagaaattagaggtgaacaatgaacactgttaattta  
cagaaatgtatcccactccagtataacagctttctgtgaggctatctcctccagactgtggctct  
gggaggggtggggcctgagtcagggtcctagggactagtgtgtgtcttcatttattccttgaata  
acgaaacgcttgagcatcagggactgtgctagcaccacaaaatccagtgggtgaacaacatggcttc  
atgggttctactgtctagaaagggaagcacattaaagaaaaaatcatttgcgtaattatttaata  
tacaactgtgatgggtactatcacaaagggaaggccaagagggaacctgatttagatgaggttg  
cagggaaggcctctctgaggaagcagcacttacactaagccatgaaggatgaataggagctagtc  
agctgaggtgagtttctgcgtagggaacagcatgtgcaaagggtctggggcaggagggagtggtg  
gtgtcctggaagaactgccagaagctgtgtgccccagggttcagacagtggtggaagaggggact  
acaggaggctgaggagataggcagggactggaccataaaagatctgtgggtcatgatgtgcat  
tggtctttatcctaaaagtgtgaaagttagtgaaagtttgaagcaggagagcatgtgatca  
gatctgcaatgcaaaaagaccaattcttggctcttctaggaaactgaattggagaaggccagagt  
acgtggaaatgacctgtcagtaggacattgtactgatgcagggaagagatgatgggtgctcagac  
caagatggccggccaaagacatagaggttccagggaaggcattctagattcttaggaattagggga  
gaactttgtgatacaaggaacatggggatgagaaggaagggtgtccaggttgacccccagggttact  
aacctgctcagcaggatgagagtggtccattcactaagccaggggaccctaggaggtgtggctac  
tttgaggtgtgggggagaggtccaagtgaggatgccaagcaggtaactgcctccacggacataca  
aacaaggccgtggcattgatgagatcgggtggggaaaagggttagccccaacctggaggaaat  
ctcagatgtagaggtcacatggaggagaatataggaaaggaaattgaagttagagtgtcagatgc  
aggagaaaaatcagcgcataataaccaagccaaggggagggagtgcctcaagaaggaggagagga  
agggtcaggacagccaaaactcctgagggccaagaaagacaagacctggaaaatgtcattaaattc  
aggcttatggaggctacaggtgaccttagtgagaccagtgaacagagggtggcagctggagag  
gatccatgctaataatgaaggaactatctgcaaagggtatgttcttaatttcagggatcatgtg

FIG. 18 (9 of 10)

tattgtgtgatacacgagtggtgtgctatgaacacaccttgggaaggagtggtgagaggatccttaa  
cattttacctgtgtacttttgtcttctccttttcaacagcctaaatggaaacctgataaaacca  
gaggaggccaaagtctatgaagatgagaagcggattatctgtttctgagaggatgctttcctgtt  
catgggggttttgcctggagcctcagcagcaaatgccactctgggcagtcttttgtgtcagtggt  
cttaaaggggcctgcgagggcgggactatcaggagtcactgcctccatgatgcaagccagcttc  
ctgtgcagaaggctgtgtcggcaaaactccctaagtaccgcgtacaattctgcagaaaaagaatgt  
gtcttgcgagctgttgtagttacagtaaatacactgtgaagagactttattgcctattataa



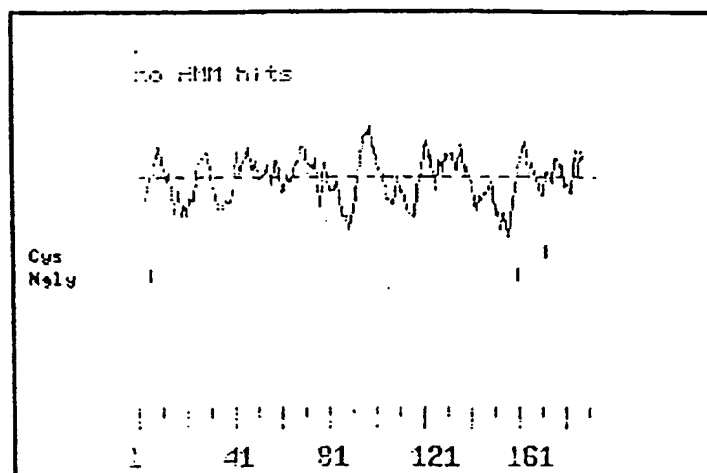


FIG 20



1 CGCGTCCGGCTGCAGCGGGGTGAGCGGCGGCAGCGCCGGGGATCCTGGAGCCATGGGGC  
CGGCAGGCCGACGTGCGCCCACTGCGCGCGTGC CGCGCCCTAGGACCTCGGTACCCCG  
1► M G

61 GCGCGCGGACGCCATCCTGGATGCGCTGGAGAACC TGACCGCCGAGGAGCTCAAGAAGT  
CGCGCGCGCTGCGGTAGGACCTACGCGACCTCTTGGA CTGGCGGCTCCTCGAGTTCTTCA  
3► R A R D A I L D A L E N L T A E E L K K

121 TCAAGCTGAAGCTGCTGTGGTGCCGCTGCGCGAGGGCTACGGCGCATCCCGCGGGGCG  
AGTTTCGACTTCGACGACGCCACGCGACGCGCTCCCGATGCGCGGTAGGGCGCCCCG  
23► F K L K L L S V P L R E G Y G R I P R G

181 CGCTGCTGTCCATGGACGCGCTTGACCTCAACGACAAGCTGGTCAGCTTCTACCTGGAGA  
GCGACGACAGGTACCTGCGGAACCTGAGGTGGCTGTTCGACCAGTCGAAGATGGACCTCT  
43► A L L S M D A L D L T D K L V S F Y L E

241 CCTACGGCGCCGAGCTCACCCTAAGCTGCTGCGCGACATGGGCTG CAGGAGATGGCCG  
GGATGCCCGGGCTCGAGTGGCGATTGCACGACGCGCTGTACCCGACGTCCTCTACCGGC  
63► T Y G A E L T A N V L R D M G L Q E M A

301 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCC  
CCGTGACGTCGCGCGGTGCGTGGTCCCGAGACCTCGGCGCGGTGACCCTAGGTCCGGG  
83► G Q L Q A A T H Q G S G A A P A G I Q A

361 CTCCTCAGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA  
GAGGAGTCAGCCGTGCGGTTCGGTCCGACGCTGAAATATCTGGTCTGGCCCGACGCGAAT  
103► P P Q S A A K P G L H F I D Q H R A A L

421 TCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTTGACGG  
AGCGCTCCCACTGTTTGCAACTACCGACGACCTACGAGACATGCCCTTCCAGGACTGCC  
123► I A R V T N V E W L L D A L Y G K V L T

481 ATGAGCAGTACCAGGCAGTGCGGGCCGAGCCACCAACCCAAGCAAGATGCGGAAGCTCT  
TACTCGTCACTGGTCCGTACGCGCGGCTCGGGTGGTTGGGTTCTTCTACGCTTCGAGA  
143► D E Q Y Q A V R A E P T N P S K M R K L

541 TCAGTTTTCACACCAGCCTGGAACCTGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGGG  
AGTCAAAGTGTGGTCCGACCTTGACCTGGACGTTCTGAAACGAGGAGTCCGGGATTCCT  
163► F S F T P A W N W T C K D L L L Q A L R

601 AGTCCAGTCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACAC  
TCAGGGTCAGGATGGACCACTCTGGACCTCGCCTCGACTCCGAGGAAGGGTCGTTGTG  
183► E S Q S Y L V E D L E R S

661 TCCGGTCAGCCCCCTGGCAATCCACCAAATCATCTGAATCTGATCTTTTTATACACAAT  
AGGCCAGTCGGGACCGTTAGGGTGGTTTAGTAGGACTTAGACTAGAAAAATATGTGTTA

721 ATACGAAAAGCCAGCTTGAA  
TATGCTTTTTCGGTCAACTT

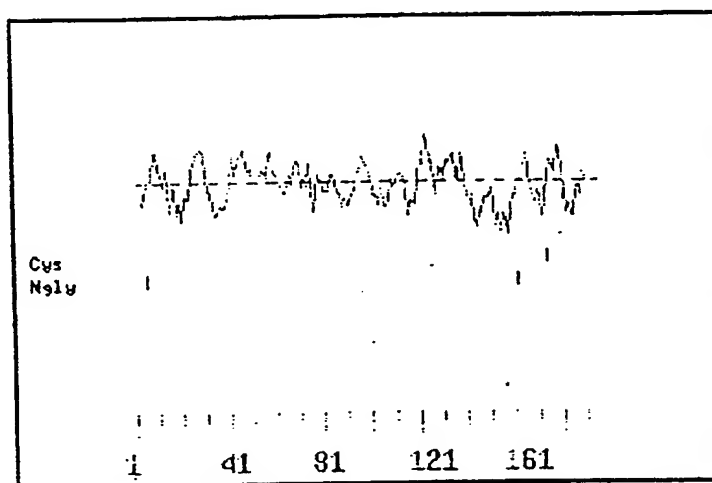


FIG. 22

40/58

ALIGN calculates a global alignment of two sequences

version 2.0uPlease cite: Myers and Miller, CABIOS (1989)

&gt; hCARD5-DNA 740 aa vs.

&gt; mCARD5-DNA 763 aa

scoring matrix: pam120.mat, gap penalties: -12/-4

68.2% identity; Global alignment score: 2377

```

      10      20      30
inputs C---GCGTCCGGCTGCAG-CGGGGTG-----AGCG-GCGGCAGC-----GGC
      :  :::::  :::  ::  :::  :::::  :::  :::  :::  :::  :::  :::
      CCACGCGTCCGGCAGCAGGCAGGCTGCAGCAGGCGAGCAGCAGCAAGAGTAAAGGTGAC
      10      20      30      40      50      60

      40      50      60      70      80      90
inputs CGGGGAT-----CCTGGAGCCATGGGGCGCGCGCGACGCCATCCTGGATGCGCTGGA
      ::  ::  :  ::  :::::  :::::  :::::  ::  ::  ::  ::  ::  ::  ::  ::
      CGCGGCTGCCACCCAGAGCCATGGGGCGGGCACGAGATGCCATCCTGGACGCTCTTGA
      70      80      90      100     110     120

      100     110     120     130     140     150
inputs GAACCTGACCGCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGCGGTGCCGTGCG
      :::  :::  :  ::  :::::  :::::  :::::  :::::  :::::  :::::
      AAACCTGTGACGGGATGAACCTCAAAAAGTTCAAGATGAAGCTGCTGACAGTGCAACTGCG
      130     140     150     160     170     180

      160     170     180     190     200     210
inputs CGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGCTGTCCATGGACGCCTTGGACCTCAC
      :::::  :::::  :::  ::  :::  :::  :::  :::  :::  :::  :::
      AGAAGGCTATGGGCGCATCCCGCGGGGCCCTGCTGCAGATGGACGCCATAGATCTCAC
      190     200     210     220     230     240

      220     230     240     250     260     270
inputs CGACAAGCTGGTCAGCTTCTACCTGGAGACCTACGGCGCGAGCTACCGCTAAC-GTGC
      :::::  :::::  :::  ::  :::  :::  :::  :::  :::  :::  :::
      TGACAAACTGTGTCAGCTACTATCTGGAGTCGTATGGCTTGGAGCTCAC-AATGACTGTGC
      250     260     270     280     290

      280     290     300     310     320     330
inputs TGC GCGACATGGGCTGCAGGAGATGGCCGGGCGAGCTGCAGGCGCCACGCACCAGGGCT
      :  :  :::::  :::  :::  :::  :::  :::  :::  :::  :::  :::
      TTAGAGACATGGGCTTACAGGAGCTGGCTGAGCAGCTGCAAACGACTAAAGA--AGAG-T
      300     310     320     330     340     350

      340     350     360     370     380     390
inputs CTGGAGCCGCGCCAGCTGGGATCCAGGCCCTCCTCAGTCGGCAGCCAGCCAGGCCTGC
      :::::  :  :::  ::  :  :  :::  :::  :::  :::  :::  :::
      CTGGAGCTGTGGCAGCTGCAGCCAGTGTCCCTGCTCAGAGTACAGCCAGAACAGG---AC
      360     370     380     390     400     410

      400     410     420     430     440     450

```

FIG. 23 (1 of 2)

```

inputs  ACTTTATAGACCAGCACCGGGCTGCGCTTATCGCGAGGGTCACAAACGTTGAGTGGCTGC
        ..... : : ..... : : ..... : : : : :
        ACTTTGTGGACCAGCACAGGCAAGCACTCATTGCCAGGGTCACAGAAGTGGACGGAGTGC
           420         430         440         450         460         470

           460         470         480         490         500         510
inputs  TGGATGCTCTGTACGGGAAGGTCCTGACGGATGAGCAGTACCAGGCAGTGGCGGGCCGAGC
        ..... : : : : : : : : : : : : : : : : : : : : :
        TGGATGCTTTGCATGGCAGTGTGCTGACTGAAGGACAGTACCAGGCAGTTTCGTGCAGAGA
           480         490         500         510         520         530

           520         530         540         550         560         570
inputs  CCACCAACCCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACCAGCCTGGAACCTGGACCT
        ..... : : : : : : : : : : : : : : : : : : : : :
        CCACCAGCCAAGACAAGATGAGGAAGCTCTTCAGCTTTGTTCCATCCTGGAACCTGACCT
           540         550         560         570         580         590

           580         590         600         610         620         630
inputs  GCAAGGACTTGCTCCTCCAGGCCCTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGG
        ..... : : : : : : : : : : : : : : : : : : : : :
        GCAAGGACTCCCTCCTCCAGGCCTTGAAGGAAATACATCCCTACTTGGTGGATGGACCTGG
           600         610         620         630         640         650

           640         650         660         670         680
inputs  AGCGGAGCTGAGGC-TCCTTCCCAGCAACACTCCGGTC-AGCCCCTGGCAAT-CCCAC-C
        ..... : : : : : : : : : : : : : : : : : : : : :
        AGCAGAGCTGAGGTATCTTTTCCAGCTACATT--ATCTAGCTCCTGACTTTGTATACAC
           660         670         680         690         700         710

           690         700         710         720         730         740
inputs  AAATCATCCTGAATCTGATCTTTTATACACAATATACGAAAAGCCAGCTTGAA
        ..... : : : : : : : : : : : : : : : : : : : : :
        AATTTTGAaaaaaCAATT-TGTATTGTGTTTAAAAAAAAAAAAAAAAAAGG
           720         730         740         750         760

```

42 / 58

ALIGN calculates a global alignment of two sequences

version 2.0uPlease cite: Myers and Miller, CABIOS (1989)

&gt; hCARD5-protein 195 aa vs.

&gt; mCARD5-protein 193 aa

scoring matrix: pam120.mat, gap penalties: -12/-4

71.8% identity; Global alignment score: 712

```

      10      20      30      40      50      60
inputs MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY
      .....
      MGRARDAILDALENLSGDELKKFKMLLTQVLRREGYGRIPRGALLQMDAIDLTDKLVSY
      10      20      30      40      50      60

      70      80      90      100     110     120
inputs LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA
      .....
      LESYGLELTMTVLRDMGLQELAEQLQT-TKEESGAVAAAASVPAQSTARTG-HFVDQHRQ
      70      80      90      100     110

      130     140     150     160     170     180
inputs ALIARVTNVEWLLDALYGKVLTDQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA
      .....
      ALIARVTEVDGVLDALHGSLTEGQYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLQA
      120     130     140     150     160     170

      190
inputs LRESQSYLVEDLERS
      .....
      LKEIHFPYLVMDLEQS
      180     190

```

FIG. 24

43/58

1 CCGCGTCCGGACTTCCCTTCCAGTGTGTGTTCCCTCTCTGCTCTCTCCAAACAGAAAGGTATTTTG  
GGGCGCAGGCTGAAGGAAGGTCACAAACAAGGAGAGACGAGAGAGGTTGCTTCCATAAAAAAC

66 GCATGTTTATCTTTGCTAAGTAGGATTTCTGTCTTCTTTGTTAACACAGATTCTTTCTGTGC  
CGTACAAATAGAAACGATTTCCTTAAAGACAGAAAGAAACAATTGTGTCTAAAGAAAGACACG

131 CAGAATGACCTGATCCATTTCTGGTGTGTAGAAAGCCATGGCTTCAGAGGGTGCTTCCCTCAGAA  
GTCTTACTGGACTAGGTAAAGGACCAACATCTTTCCGTACCGAAGTCTCCACGAAGGAGTCTT  
1 M A S E G A S S E

196 ATCATAGAAAAACAGCGAACAAAGTTGCTCAGTGTCTTCCAACAAGATCCCGACTCTATCTTGG  
TAGTATCTTTTGTGCGCTTGTTCACAGGAGTCAAGGAGGTTGTCTAGGGCTGAGATAGAACCT  
10 I I E K Q R T K L L S V L Q Q D P D S I L D

261 CACGTTAACCTCTCGGAGACTGATTCTGAGGAGGAGTATGAGACTCTAGAGGCAATTACAGATC  
GTGCAATTGGAGAGCCTCTGACTAAAGACTCCTCCTCATACTCTGAGATCTCCGTAAATGTCTAG  
31 T L T S R R L I S E E E Y E T L E A I T D

326 CTCTGAAGAAAAGCCGGAAGCTGTAAATTTTGATCCAGAAGAAGGGAGAGGACAGCTGTTGTGT  
GAGACTTCTTTTGGGCTTCGACAAATAAACTAGGTCTTCTTCCCTCTCTGTGACAAACAAC  
53 P L K K S R K L L I L I Q K K G E D S C C C

391 TTCTCAAGTGTCTGTCTAATGCCTTTCCACAGTCAGCTTCCACCTTGGGTTTAAAGCAGGAAGT  
AAGGAGTTCACAGACAGATTACGGAAAGGTGTGAGTCGAAGGTGGAACCCAAATTTCTGCTCTCA  
75 F L K C L S N A F P Q S A S T L G L K Q E V

456 TCCACGGCAGGGACTGGAGAGGTTGTGAGGTGAGCAGGGGTTTGAAGATCCCTTTTCTCTTG  
AGGTGCGCTCCCTGACCTCTCCAAACAGCTCCACTCGTCCCCAAACCTTCTAGGGAAAAGAGAAC  
96 P R O G T G E V V E V S R G L E D P F S L

521 GGACCATAACCCAGAAATAGCAGAGCTCTCAGAAGAGAAAGAATGCCGGGTCTGGGAGCTCCG  
CCTGGTATTGGGCTCTTATCGTCTCGAGAGTCTTCTCTTCTTACGGGCCAGACCTCGAGGC  
118 G T I T P E I A E L S E E K E C P G L G A P

586 GAGTCTTCACTGCAAGGAAGCAGCCACAGGGAACCGGAAGTACCTTCTTGGGAGAAACAGGA  
CTCAAGAAAGTGGACGTTCTTTCTGCGGTGTCCCTTGGCCTTCATGGAAGAACCCTCTTAGTCTCT  
140 E F F T C K E S S H R E P E V P S W E N O E

651 AGGCGTGGTGACAGCAAGTCACCGCTCCGCTTCACTCAAAGGAGTTGAGTATGAAGTTCCAG  
TCCCGCACCACTGTGCTTCACTGGCGAGGCGCAAGTCAGTTTCTCAACTCATACTTCAAGGT  
161 G R G A Q Q V T A P R S V K G V E Y E V P

FIG. 25 (10F7)

44/58

716 CAAGTATCTCCCTCTTAAGCGACGGGCAGAGATACGAGGAGCCAGATGATTCGCTGTACTTAGAA  
GTTCATAGAGGGAGAATTCCGCTGCCCGTCTCTATGCTCCCTCGGTCTACTAAGCGACATGAATCTT  
183▶ A S I S L L S D G O R Y E E P D D S L Y L E

781 GAAGGGGAAGGTGAAGAGTCTCTTGGGTACCCCTGAAGATGTTTGGAGGAAGGGCCGCGATGA  
CTTCCCTTCCACTTCTCAGAGAACCCATGGGACTTCTACAAAACCTCTTCCCGGGCGCTACT  
205▶ E G E G E E S L G Y P E D V L E E G A G D D

846 CCCACAGTGCCTTTGTATATGATAGTGAGGAGGAATGCGAGTATGAGGAAAACATGGGCTCCTCCG  
GGGTGTACGAAACATATACTATCACTCCTCTACGCTCATACTCCTTTTGTACCCGAGGAGGC  
226▶ P Q C F V Y D S E E E C E Y E E N M G S S

911 GTGAAGACAGTAGCTGCGACGACACTTCAGAGACCTGCGTTCCATTGGAAGGGGAGAAAAGCGCT  
CACTTCTGTCACTCGACGCTGCTGTGAAGTCTCTGGACGCAAGGTAACCTTCCCTCTTTTCGCGA  
248▶ G E D S S C D D T S E T C V P L E G E K S A

976 GAAGAAAGAAAAAGAGTGTTCACACGTCCTGTCTGTGTTGAACATGGATAGAAACAGAAAGCT  
CTTCTTTCTTTTCTCACAAGTTGTGCAGGACAGGACAAACTTGTACCTATCTTTGTCTTTTGA  
270▶ E E R K R V F Q H V L S C L N M D R N R K L

1041 TCTCCAGAGTTCGTGAGGCAGTTTCCATAGACCGAGGATGTGAGTGGACACCCAGACCCAG  
AGAGGGTCTCAAGCACTCCGTCAAAAGGTATCTGGCTCCTACACTCACCCTGTGGGTCTGGGGT  
291▶ L P E F V R O F S I D R G C E W T P K T P

1106 GAGACTTAGCTTGAATTTCTTGATGAAAGTTTCAGGCTTTAGACTCGACAGCCAGAGATTCTATC  
CTCTGAATCGAACCTTAAAGAACTACTTTCAAGTCCGAAATCTGAGCTGTGCGTCTCTAAGATAG  
313▶ G D L A W N F L M K V O A L D S T A R D S I

1171 CTGAGGCCCCGAGGTGCGGGGTGAAGAGAATGAAGAATGCGCGCTGGAATAGAGAAGTTAGGCAT  
GACTCCGGGCTCCACGCGCCACTTCTCTTACTTCTTAACGCGCGACCTTATCTCTTCAATCCGTA  
335▶ L R P E V A G E E N E E L P A G I E K L G I

1236 TGGAGACCCCAAACCATCCATCCCTGGATGTCTCTGCGCTGCATGCTTTGTGCAGACAGCT  
ACCTCTGGGGTTTGGTAGGTAGGGGACCTACAGGAGACGCGGACGTACGAAACAOGTCTGTGGA  
356▶ G D P Q T I H P L D V L C A C M L C A D S

1301 CCTTGCAGCGTGAAGTCATGTCAAACATGTACCAATGCCAGTTTGTCTTCCCTGCTACTGCCA  
GGAACGTCGCACTTCAGTACAGTTTGTACATGGTTACGGTCAAACGAGAAGGGGACGATACGGT  
378▶ S L O R E V M S N M Y O C O F A L P L L L P

1366 GATGCTGAGAACAACAAAACCTCTTAATGGTAGGGGCCATGAAGGACTTAAAGCAGCCCTCAGC  
CTACGACTCTTGTGTTTGTGAGAATTACCATCCCGGTACTTCTGAATTTGTGCGGAGTCTG  
400▶ D A E N N K N L M V G A M K D L K Q P S A

FIG. 25 (207)

45/58

1431 ACAGTCTCAGGAGGGCCCCCTCAGGGAAACAGACACATTTCTGGGTCTCACAAGATGCCTGTCA  
TGTCAGGAGTCTCTCCCGGGGAGTCCCTTTGTCTGTGTAAGACCCAGAGTGTCTACGGACAGT  
421▶ O S S G G P L R E T D T F L G L T K M P V

1496 TCTCTTTTGTGCGACTAGGACGCTGCAGCTTCTCCAAGTCCAGAATTGTTAACACACTGCTCAGC  
AGAGAAAACACGCTGATCCTGCGACGTCGAAGAGGTTCAAGGTTTAAACAATTGTGTGACGAGTCG  
443▶ J S F V R L G R C S F S K S R I V N T L L S

1561 TCCTCCAGCAGAAACCATACCCGATTTTCTCCATCAGGATCTGTCTGTCCCTGTGCTTCCTCG  
AGGAGGGTCTGCTTTTGGTATGGGCTAAAAGGAGGTAGTCTAGACAGACAGGGACACGAAGGAGC  
465▶ S S Q Q K P Y P I F L H Q D L S V P V L P R

1626 GCAAATTTCTGACGGCCTGGTGGAGTGACATGGTGCTTTCTCTGACAAGTTGCTGAAGGAAAGCC  
CGTTTAAAGACTGCCCGACCACCTTCACCTGTACCACGAAAGGACTGTTCAACGACTTCCTTTCCG  
486▶ Q I S D G L V E V T W C F P D K L L K E S

1691 CGCATGCTTTCCAGAAACCTGTGTGTGGCCAACTTCGTGGAGATTAGAAAGCTTTTGGATA  
GGGTACGAAAGGTCTTTGGACAACGACACCGGTTGAAGCACCTCTAAATCTTTCGAAAACCTAT  
508▶ P H A F O K P V A V A N L R G D L E S F W I

1756 CAATTTGGTTTCTGGTAGAAGTTTCTCCGGTCTTTCTTTTTCACAGACTGCCTTGGTGAGAA  
GTTAAACCAAAGGACCATCTTCAAAGGAGGCCAGAAAAGAAAAGTGTCTGACGGAACCACTCTT  
530▶ Q F G F L V E V S S G L F F F T D C L G E K

1821 GGAATGGGACTTGCTAATGTTTTAGGAGAGGACACCATTTGAACGGTGCTACTTTATCCTCAGTC  
CCTTACCCTGAACGATTACAAAAATCCTCTCTGTGGTAACTTGCCACGATGAAATAGGAGTCAG  
551▶ E W D L L M F L G E D T I E R C Y F I L S

1886 CCCAGGCTAAGGAGAGTGAAGAAGCCAGATTTTCCAAAGGATCTTAAACTGAAGCCATCTCAG  
GGGTCCGATTCTCTCACTTCTTCCGGTCTAAAAGGTTTCTTAGGATTTTGACTTCGGTAGAGTC  
573▶ P Q A K E S E E A O I F Q R I L K L K P S Q

1951 CTAAGTTTGGGAAGCTGAGGAAGCTGGGGATAGAAGGAAGACTATGGAGGCCCTTCAAGCTGC  
GATGACAAAACCTTCGACTCTTCGACCCCTATCTTCTTCTGATACCTCCGGAAGTTTCGACG  
595▶ L L F W E A E E A G D R R K T M E A L Q A A

2016 CCTCCAGGAAGTAATGTCTCTCCACTCAGATGTGTGTCCCTTGAAGAGATGGCCTCTCTGGCCA  
GGAGGTCTTTCATTACAGGAGAGGTGAGTCTACACACAGGGAACCTCTCTACCGGAGAGACCGGT  
616▶ L Q E V M S S P L R C V S L E E M A S L A

2081 GGGAGCTGGGCATTTCAGGTAGACCAAGACTTTGAAGTTACTCAAGATATTCAAGTTTCCCCACA  
CCCTCCACCCGTAAGTCCATCTGGTTCTGAAACTTCAATGAGTTCTATAAGTTCAAAGGGGTGT  
638▶ R E L G I Q V D Q D F E V T Q D I Q V S P T



46/58

2146 ACAGTTGAAGGTGAAAACCAACAACCATGTAGTCAGACCAAAAGCCCGGCTGAAAGCGGAGCTCA  
TGTCAACTTCCACTTTTGGTTGTGGTACATCAGTCTGGTTTTCGGGCCGACTTTCGCCTCGAGT  
660 T V E G E N Q Q P C S O T K S P A E S G A Q  
-----

2211 GGAGCCAATCAGAGAGCCAGGGGCTCAATGTGACGACAGTCAGAATGCTCCGGTTTTCATCAGA  
CCTCGGTTAGTCTCTCGGTCCCCGAGTTACACTGCTGTCTAGTCTTACGAGGCCAAAAGGTAGTCT  
681 E P I R E P G A Q C D D S Q N A P V F H Q  
-----

2276 CTCCAGTATACATGCCTTATCCAGCACACCCATGGGCTTTGGCCATCAAAGCTGGAGGTAACCTT  
GAGGTCATATGTACGGAATAGGTCTGTGGGTACCCGAAACCGGTAGTTTCGACCTCCATTGAAA  
703 T P V Y M P Y P A H P W A L A I K A G G N F  
-----

2341 TACCACGTTCCCTTTGAATGCCCCCTGGTTATGGGCTCCCACTTTGGATCAGCAGAGGGCTAAG  
ATGGTGCAAGGAACTTACGGGGACCAATACCGAGGGTGAAACCTAGTGTCTCTCCGATTCT  
725 Y H V P L N A P W L W A P T L D H S R G L S  
-----

2406 TGGTTCTTTCCATTCCCATGCTAAACCCACTCACTCTAAGGCCTTCCAAGCTAACTGCCACCATC  
ACCAAGAAAGGTAAGGGTACGATTGCGGTGAGTGAGATTCCGGAAGGTTGATTGACGGTGGTAG  
746 G S F H S H A K P T H S K A F Q A N C H H  
-----

2471 CCCATCCCTCCCATGCTAAACCCACTCACTGTGAATCCCTCTCATGCTAAACCCACTCATGTGCAG  
GGGTAGGGAGGGTACGATTGCGGTGAGTACACTTAGGGAGAGTACGATTGGGGTGAGTACACGTC  
768 P H P S H A K P T H V N P S H A N P T H V Q  
-----

2536 CCTTGCATGCTAAACCCACTCACTCTAAGGCCTTCCAAGCTAAACCCACTCCCTCTCAGACCTCT  
GGAACGTACGATTGCGGTGAGTGAGATTCCGGAAGGTTGATTGCGGTGAGGAGAGTCTGGAGA  
790 P C M L N P L T L R P S K L N P L P L R P L  
-----

2601 TGGAGCCAAGCTAACTGCAATCATGCCCATCCCTCCCTTGCTAAACCCCTCTCATACGAATCCCTC  
ACCTCGGTTTCGATTGACGTTAGTACGGGTAGGGAGGGAACGATTGCGGAGATATGCTTAGGGAG  
811 G A K L T A I M P I P P L L N P L I R I P  
-----

2666 TGATGCTAAACCCACTCACTGTGAGCCTTCCCATGCTAAACCCGCTCATCTACAGTCTTCCCAA  
ACTACGATTGCGGTGAGTACACGTCGGAAGGTTACGATTGCGGAGTAGATGTCAGAAGGGTTT  
833 L M L T P L M C S L P M L N P L I Y S L P K  
-----

2731 CAAAACCCCTCCCATCCCAATCTACTGCAGTTCACGGCACACAAACCTCAGCAGTCCAGTCTAA  
GTTTTGGGAGGGGTAGGGTTAGATGACGTCAAGTGCCGTGTGTTGGAGTCTGAGGGTCAGATT  
855 Q N P P H P N L L O F T A H K P Q Q S Q S K  
-----

2796 GCCTTCTCAGCAGAGACCCAGTCAGCCTAAATCATTCAGACCAAGCCTTCACAGGCCAGGGCCT  
CGGAAGAGTCGTCTCTGGGTCACTCGGATTAGTAAGGTCTGGTTCGGAAGTGTCCGGTCCCGGA  
876 P S Q Q R P S Q P K S F Q T K P S Q A R A  
-----

47/58

2861 GCCACCCAAGAGCAGGAGACGTTAAGAACATACTCTGGAGATCTGGGAAATAAAGTATGGGCT  
CGGTGGGTTCCTCGTCCCTCTGCAATTTCTTGTATGAGACCTCTAGACCCCTTATTTTCATACCCGA

898 C H P R A G R R

2926 TTGCTTAAGTATTCCTTTTTCATATAGCAAGCTGAAGAAAAGTTTGTAGTAAAGACTGATAAAAGT  
AACGAATTCATAAGAAAAAGTATATCGTTCGACTTCCTTTCAAAATCACTTTCTGACTATTTTCA

2991 AGCAAAACCCAAAAAAGGTATGCAAAGTCTTAAGTGCATAGCAAAGTATCCAAGTGTGGGAAATA  
TCGTTTTGGGTTCCTTCATACGTTTCAGAAATTCAGTATCGTTTCATAGGTTTACACCCCTTTAT

3056 TGGAAGCAGTTAAAAGTAGAATCTGGCTGGGCATGGTGGCACACATCTACAGGGTTTAGCATGGG  
ACCTTCGTCAATTTTCATCTTAGACCGACCCGTACCAACCGTGTGTAGATGTCCCAAATCGTACCC

3121 AGGGCTCTGTTCATCCCAACTCAGAGAAGCAGGCAGATCTCTGTGTGTTTGAGGCCAGTCTGGTCT  
TCCCGAGACAGTAGGGTTGAGTCTCTTCGTCCGTCTAGAGACACACAACTCCGGTCAGACCAGA

3186 ACATAACAACGACACAAGCAAGTCTTACATCAGCCATACTACAAAATGAGACCCCATCTGGGGAC  
TGTAATTGTGCTGTGTGTTTCAGGATGTAGTCCGTATGATGTTTACTCTGGGGTAGACCCCTG

3251 AAAAGGGTTGGATCTAACATCAAACCAAGAAATCAGTCAAGTATTCAGAAGGCATCATTAAAT  
TTTTCCTCAACCTAGATTGTAGTTTGGTTTCCTTAGTCAAGTTCATAAGGTCTTCGTAGTAATTAA

3316 ACACTCAGTGGGTACCACAACCAACCATACTCGACAACCTAACCCCTAAAGGAGCAAGAAGGA  
TGTGAGTCAACCAATGTTGTTGGTTTGGTATGAGCTGTTGATTGGGGATTTCCTCGTTCCTCTCT

3381 GTTGGGTGGGTGTAGGCTGAACATGATTGGGGAAGAACTGAAGATAGATAAGGTCAATTCGTAAT  
CAACCAACCCACAATCCGACTTGTACTAACCCTTCTTGACTTCTATCTATTCCAGTAAGCATTA

3446 ACAGGTTATGGGACTTGTCAAATCCATTAAATGCAATATTAAGAAGCAGTGGGAATCTTAAGGCT  
TGTCGAATACCCGAAACAGTTTAGGTAATTTACGTTATAATTTCTTCGTACCCCTTAGAATTCGA

3511 ACATTAAGCTCCAGTGAGTCGCAACCCCTCCCTATTAGATGATGTGAGATTTGAACCCCAAGTGAA  
TGTAATTCGAGGTCACTCAGGTTGGGAGGGGATAATCTACTACACTCTAAACTTGGGGTCACTT

3576 TGGGGTGTCTGATAGCCCGTGTGTGTGACAACTGTGTAATTATAAAGTGATGAAAACGTGGG  
ACCCACACAGACTATCGGGCACACACTGTTTGACACATTAAATTTTCACTACTTTTGACCCC

3641 AGTTCAGCTTATCTGRTTGAAGAAAGGCTGCTTCAGAGGTGCCTTGGTTTTGGGTTTATGATCA  
TCAAGTCGAATAGACACAACCTCTTTCGACGAAGTCTCCACGGAAACCAACCCAAATACTAGT

FIG 25. (5 of 7)

48 | 58

3706 GCCACTGAGCAGATACTCTGCACCATTTGGTACAGTTAAATCAGCTTGCTTCTGGTAATAGCCCCA  
CGGTGACTCGTCTATGAGACGTGGTAACCATGTCAATTTAGTCCAACGAAGACCAITATCGGGGT  
-----

3771 ATCTACCACATTTATCCCTTACAGGCGGAAATAATGAATGATCAGCAAAACATCCAATTTTACCT  
TAGATGGTGTAAATAGGGAATGTCCGCTTTTATTACTTACTAGTCGTTTGTAGGTTAAATCGA  
-----

3836 TAACCTTGGTACTGATTTGTATATGTATCATTCTTTATATAATAGCTAAGAAAAATTAGCTCATT  
ATTGGAACCATGACTAAACATATACATAGTAAGAAATATATTTCGATTCCTTTAAATCGAGTAA  
-----

3901 AGGGGTCTGATATATTAGTTTAAATGGTTTGAAGTCAGAAATGTGTTAGTTTAAATTTTAGAGT  
TCCCCAAGACTATATAATCAAATTACCAAACCTTCAGTCTTTACACAATCAAAAATTAAATCTCA  
-----

3966 TAATTGAAAAATATTGAGATGAATTTACAAAGGCTATAAGTAATGTTTGAGAGGGTTATAATTTTT  
ATTAACCTTTTATAACTCTACTTAAATGTTTCGATATTTCATTACAACTCTCCCAATATTAAAAA  
-----

4031 GTAGACTCATACTGTTCTGAACATTTGGATAGCTTCTCGTAGTTAGCAGTGTTATAGAAGAATAT  
CATCTGAGTATGACAAGACTGTAAACCTATCGAAGAGCATCAATCGTCACAAATATCTTCTTATA  
-----

4096 ATTTGATTTCAGGTATTTAACCAGAGCTGCTCTTAGTTTAAAGTGTACCAAGAGTCAATAAAAG  
TAAACTAAGTCCATAAATGGTCTCGACGAGAATCAAAAATTCACAGTGGTTCAGTTATTTTC  
-----

4161 GCTACATTATCTGAACATGTGGGAACACAACCTGTGACCTTACACTTAAGAGACTGAGGAAGGGAA  
CGATGTAATAGACTTGTACACCTTGTGTTGACACTGGAAATGTGAATTCCTGACTCCTTCCCTT  
-----

4226 ATCAAGGTTCAAGCCAGCAGCACATAGTGAGACCAGGTCTCAAGACACAAAACTATCCACCTTA  
TAGTTCCAAGTTCGGTCTGCTGTATCACTCTGGTCCAAGTTCGTGTGTTTGTATAGGTGGAAT  
-----

4291 AGGAAGATTTTAAAAATTGCCCTCATTAAAGAAATAAGTAAGATTATAAATTGGACTAAATGTCA  
TCCTTCTAAAAATTTTAAACGGAGTAATTCCTTATTTCATTCTAAATATTTAACCTGATTTACAGT  
-----

4356 CATCTTTGAACCTTATGACTGTTTAAATTTTGAAGTTAAATTTTATTATGTATGCGTGT  
GTAGAACTTGAATACTGACAAATAAAAAACTGAATTTCAAATTAATAATACATAACGACA  
-----

4421 GTTGTATGTGTGTGCACATGTGTGCCACTGCATGTATGTGGAGGCCATCAGACAATGTTGTAGAG  
CAACATACACACAGTGTACACACGGTGACGTACATACACCTCCGGTAGTCTGTACAAACATCTC  
-----

4486 TCTGTCTTTTCCCTCTTAGCCCTATGTGTTTACCCACTGAGCTAGGCCACCTACTCCTATAAGTC  
AGACAAGAAAGGAGAAATCGGGATACACAAAAATGGGTGACTCGATCCGGTGGATGAGGATATTTCAG  
-----

49/58

4551 TAATTTTAAATAGTAAAATAGTTCTAAGAAGTCAATCAGGGAAAAAATGGCTGTCAAAGTCTCA  
ATTAAAAATTATCATTATATCAAGATCTTCAGTTAGTCCCTTTTITACCGACAGTTTCAGAGT  
-----

4616 AAGAAAAATCGTATTAGCCATGGATAGAGACTCACCTCTTGAATCATTTGTGTCTGAGAATAGCC  
TTCTTTTAGCATAATCGGTACCTATCTCTGAGTGGAGAACTTAGTAAACACAGACTCTTATCGG  
-----

4681 TAATATCACAATAATGTGTTTGTACATGTGTTAGTTAATATTGTTTTTCAGAGTATTTAATCTCTC  
ATTATAGTGTATTACACAAACATGTACACAATCAATTATAACAAAAGTCTCATAAATTAGAGAG  
-----

4746 ATGATTATTGTAAAGATGAAAAAAGAAATAGTGGGCAATGTATGTGAGTATTTAATTTTGCCTGA  
TACTAATAACATTCTACTTTTTTCTTTATCACCCGTTACATACACTCATAAATTAAACGGACT  
-----

4811 CAATTCTGTCTTTTAGAATGATAAATGTAAGAAGTAAAAATAAAACGGTTTCAATTCAGAACAACT  
GTTAAGACAGAAAATCTTACTATTTACATTCTTCATTTTATTTTGCCAAGTAAGAGTCTTGTGTA  
-----

4876 AAGCCAGCTCACTTAAGTCTGGGCCCTGCTGGCATTGGCTAGTCTAGCTACCCCAACCCAAACAC  
TTCCGTCGAGTGAATTCAGACCCGGGACGACCGTAACCGATCAGATCGATGGGGTGGGTMTGTG  
-----

4941 AAAAGTTTAGAGAAGAAAATGACTGAGTCAAGCTTGCCTAATGACTTTTGGACATAAAGTTTATG  
TTTTCAAATCTCTTCTTTTACTGACTCAGTTCGAACGGATTACTGAAAACCTGTATTTCAAATAC  
-----

5006 GTCCTAGAAAGCCTTAAATAAGTAGGATATAAAACATGTAAATTAACCCACACATTATGTGGGT  
CAGGATCTTTTCGAATTTTATTCATCCTATATTTGTACATTTAATTGGGTGTGTAATACACCCA  
-----

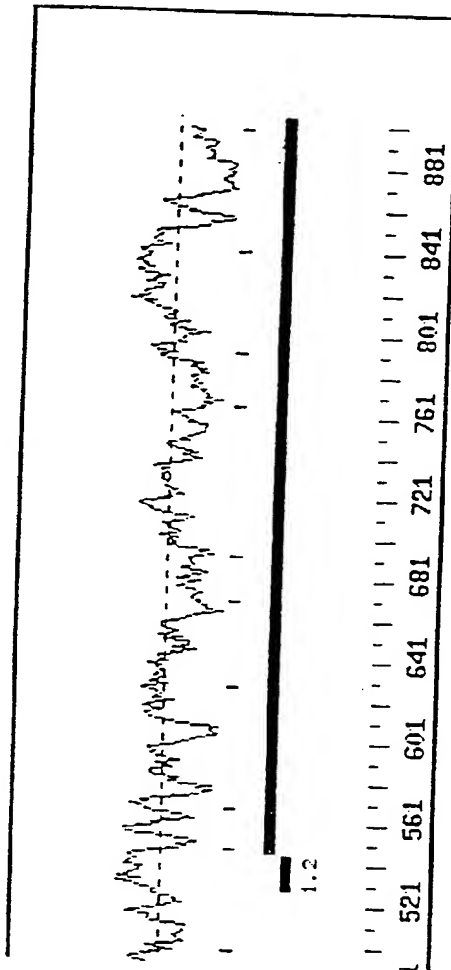
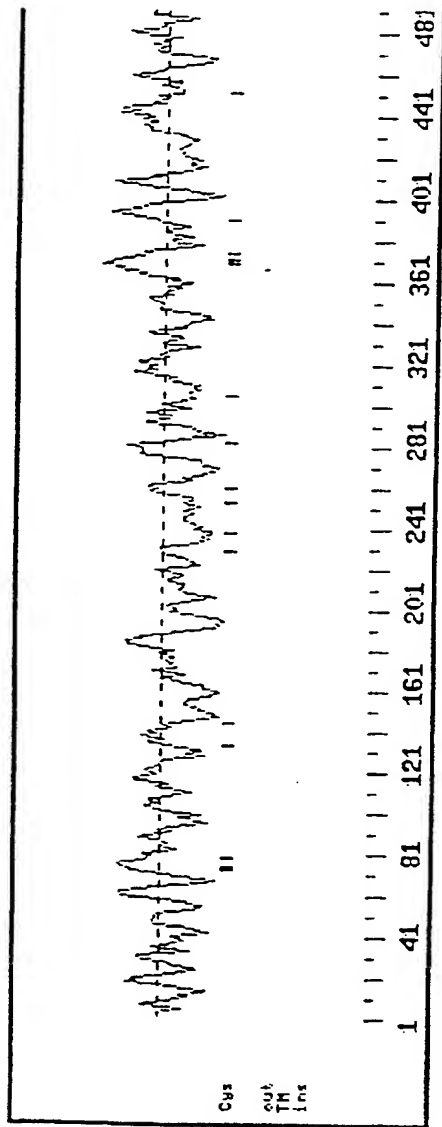
5071 TGAGAAGCAGAAAAATGTCAGTAGAACACTCGGCCAGTGCATAAAGAAGGAAGAGACCTCTGTTC  
ACTCTTCGTCTTTTACAGTCATCTGTGAGCCGGTCAGTATTCTTCCTTCTCTGGAGACAAG  
-----

5136 TGGGTTATAAACTGCTCTTTGTGCTCAATTTGTCCCTGCTTTTGTGTTGCCAGAAATGTACAAGA  
ACCCAATATTTTGACGAGAAACACGAGTTAAACAGGGGACGAAAACAAACGGTCTTACATGTTCT  
-----

5201 TTATAAAATAAACTCACTTTTACTTTTAAAAAATAAAAAAAAAAAGGGCGG  
AATATTTTATTTGAGTGAAAAATGAAAAATTTTITTTTTTTTTTTTCCCGCC  
-----

FIG. 25 (7 of 7)

50/58  
F16.26



51/58

- - - - G X H F X D Q H R X A L I A R V T X V X X V L D A L Y - G X V L T E X Q Majority																																									
1	M	E	A	R	D	K	O	V	L	R	S	L	F	L	E	L	G	A	E	V	L	V	E	G	L	V	L	O	Y	L	Y	Q	E	G	I	L	T	F	N	H	RAIID
1	-	-	-	-	G	L	H	F	I	D	Q	H	F	A	A	L	I	A	R	V	T	N	V	E	W	L	L	D	A	L	-	G	R	V	L	T	D	E	G	hCARDS	
1	-	-	-	-	G	-	H	-	V	D	O	H	F	Q	A	L	I	A	R	V	T	E	V	D	G	V	L	D	A	L	H	-	G	S	V	L	T	E	G	mCARDS	
Y Q A V R A E T T X X X K M R K L F S F X P S W N X T C K D X L L Q A L X E X X Majority																																									
41	I	C	E	I	N	A	O	T	T	G	L	R	N	T	M	L	L	D	I	L	P	S	R	G	P	K	A	F	D	T	F	L	D	S	L	Q	E	F	-	RAIID	
36	Y	O	A	V	R	A	L	P	T	N	P	S	K	M	R	K	L	F	S	F	T	F	A	W	N	W	T	C	K	D	L	L	L	Q	A	L	R	E	S	Q	hCARDS
35	Y	O	A	V	R	A	E	T	T	S	O	D	K	M	R	K	L	F	S	F	V	P	S	W	N	L	T	C	K	E	S	L	L	O	A	I	K	E	I	H	mCARDS
P Y L V E D L E X S - - - - Majority																																									
80	P	W	V	R	E	K	L	K	K	A	R	E	E	A	M																							RAIID			
76	S	Y	L	V	E	D	L	E	R	S																												hCARDS			
75	P	Y	L	V	M	D	L	F	Q	S																											mCARDS				

FIG 27

52/58

CACGCGTCCGCCGATCAGAGAGTGCTCCGAGCTGGGTTGCCCACTGTGCTTGATCTGCACTCTCCAACACTAGGC 79  
 ATCATTGACATGTTAAAGCTTAGCCAAATAGAATTGTTCTTTGTCTATTCTTTTTTAACTTTTACTTTATTTCATTAGGAT 158  
 M A T E S T P S E 9  
 GATTTTCATAATATATTTCTGTTTAGAGGAAACAGGAACA ATG GCT ACC GAG AGT ACT CCC TCA GAG 226  
 I I E R E R K K L L E I L Q H D P D S I 29  
 ATC ATA GAA AGA GAA AGA AAA AAG TTG CTT GAA ATC CTT CAA CAT GAT CCT GAT TCT ATC 286  
 L D T L T S R R L I S E E E Y E T L E N 49  
 TTA GAC ACG TTA ACT TCT CGG AGG CTG ATT TCT GAG GAA GAG TAT GAG ACT CTG GAG AAT 346  
 V T D L L K K S R K L L I L V Q K K G E 69  
 GTT ACA GAT CTC CTG AAG AAA AGT CGG AAG CTG TTA ATT TTG GTA CAG AAA AAG GGA GAG 406  
 A T C Q H F L K C L F S T F P Q L A A I 89  
 GCG ACC TGT CAG CAT TTT CTC AAG TGT TTA TTT AGT ACT TTT CCA CAG TTA GCT GCC ATT 466  
 C G L R H E V L K H E N T V P P Q S M G 109  
 TGC GGC TTA AGG CAT GAA GTT TTA AAA CAT GAG AAT ACA GTA CCT CCT CAA TCT ATG GGG 526  
 A S S N S E D A F S P G I K Q P E A P E 129  
 GCA AGC AGT AAT TCA GAA GAT GCT TTT TCT CCT GGA ATA AAA CAG CCT GAA GCC CCT GAG 586  
 I T V F F S E K E H L D L E T S E F F R 149  
 ATC ACA GTG TTC TTC AGT GAG AAG GAA CAC TTG GAT TTG GAA ACC TCT GAG TTT TTC AGG 646  
 D K K T S Y R E T A L S A R K N E K E Y 169  
 GAC AAG AAA ACT AGT TAT AGG GAA ACA GCT TTG TCT GCC AGG AAG AAT GAG AAG GAA TAT 706  
 D T P E V T L S Y S V E K V G C E V P A 189  
 GAC ACA CCA GAA GTC ACA TTA TCA TAT TCA GTT GAG AAA GTT GGA TGT GAA GTT CCA GCA 766  
 T I T Y I K D G Q R Y E E L D D S L Y L 209  
 ACT ATT ACA TAT ATA AAA GAT GGA CAG AGA TAT GAG GAG CTA GAT GAT TCT TTA TAC TTA 826  
 G K E E Y L G S V D T P E D A E A T V E 229  
 GGA AAA GAG GAA TAT CTA GGA TCT GTT GAC ACC CCT GAA GAT GCA GAA GCC ACT GTG GAA 886  
 E E V Y D D P E H V G Y D G E E D F E N 249  
 GAG GAG GTT TAT GAT GAC CCA GAG CAC GTT GGA TAT GAT GGT GAA GAG GAC TTC GAG AAT 946  
 S E T T E F S G E E P S Y E G S E T S L 269  
 TCA GAA ACC ACA GAG TTC TCT GGT GAA GAA CCA AGT TAT GAG GGA TCA GAA ACC AGC CTT 1006  
 S L E E E Q E K S I E E R K K V F K D V 289  
 TCA TTG GAG GAG GAA CAG GAG AAA AGT ATA GAA GAA AGA AAA AAG GTG TTT AAA GAT GTC 1066  
 L L C L N M D R S R K V L P D F V K Q F 309  
 CTG TTA TGT TTG AAC ATG GAT AGA AGC AGA AAG GTT CTG CCA GAT TTT GTT AAA CAA TTC 1126  
 S L D R G C X W T P E S P G D L A W N F 329  
 TCC TTA GAT CGA GGA TGT AAG TGG ACC CCT GAG AGT CCA GGA GAC TTA GCC TGG AAT TTC 1186  
 L M K V Q A R D V T A R D S I L S H K V 349  
 CTG ATG AAA GTT CAA GCA CGA GAT GTG ACG GCT AGG GAT TCA ATC CTC AGT CAC AAG GTT 1246  
 L D E D S K E D L L A G V E N L E I R D 369  
 CTG GAT GAA GAT AGC AAG GAG GAT TTG CTG GCT GGA GTG GAG AAT TTG GAA ATT CGA GAC 1306

FIG. 28 (1 of 4)

53/58

I Q T I N P L D V L C A T M L C S D S S 389  
ATA CAA ACC ATT AAT CCC CTT GAC GTG CTT TGT GCC ACC ATG CTG TGT TCA GAT AGC TCT 1366

L Q R Q V M S N M Y Q C Q F A L P L L L 409  
TTG CAA CGC CAA GTC ATG TCA AAC ATG TAT CAG TGC CAG TTT GCT CTT CCC CTG CTA CTG 1426

P D A E N N K S I L M L G A M K D I V K 429  
CCA GAT GCA GAA AAC AAC AAA AGC ATC TTA ATG CTG GGG GCC ATG AAA GAC ATT GTG AAG 1486

K O S T Q F S G G P T E D T E K F L T L 449  
AAG CAG TCA ACA CAG TTT TCA GGG GGG CCT ACA GAG GAT ACA GAA AAG TTT CTG ACT CTC 1546

M K M P V I S F V R L G Y C S F S K S R 469  
ATG AAG ATG CCT GTC ATC TCT TTT GTG CGT CTA GGA TAC TGT AGC TTC TCT AAG TCC AGA 1606

I L N T L L S P A Q L K L H K I F L H Q 489  
ATC CTC AAC ACA CTT CTC AGC CCT GCC CAG TTG AAA TTA CAC AAA ATC TTT CTT CAT CAA 1666

D L P L L V L P R Q I S D G L V E I T W 509  
GAT TTG CCT CTT TTG GTG CTT CCC CGG CAA ATC TCT GAT GGC CTG GTT GAG ATA ACA TGG 1726

C F P D S D D R K E N P F F Q K P V A L 529  
TGT TTT CCT GAT AGC GAT GAT AGA AAG GAA AAC CCC TTT TTC CAA AAG CCT GTT GCT CTG 1786

A N L R G N L E S F W T Q F G F L M E V 549  
GCT AAT CTC CGT GGA AAT CTA GAA AGT TTT TGG ACT CAG TTT GGT TTT TTG ATG GAA GTT 1846

S S A V F F F T D C L G E K E W D L L M 569  
TCT TCA GCT GTG TTT TTT TTC ACT GAC TGT TTA GGT GAG AAG GAA TGG GAC TTG CTA ATG 1906

F L G E A A I E R C Y F V L S S Q A R E 589  
TTT TTA GGA GAG GCT GCC ATT GAA AGA TGC TAC TTT GTT CTC AGT TCC CAA GCC AGG GAG 1966

S E E A Q I F Q R I L N L K P A Q L L F 609  
AGT GAA GAG GCT CAA ATT TTT CAG AGG ATA CTG AAC TTG AAG CCA GCA CAG CTA CTG TTT 2026

W E R G D A G D R R K N M E G L Q A A L 629  
TGG GAG AGG GGA GAT GCT GGG GAT AGA AGG AAG AAC ATG GAG GGC CTT CAA GCT GCC CTC 2086

Q E V M F S S C L R C V S V E D M A A L 649  
CAG GAA GTG ATG TTC TCT TCT TGC CTC AGA TGT GTG TCT GTG GAG GAT ATG GCC GCC CTG 2146

A R E L G I Q V D E D F E N T Q R I Q V 669  
GCC AGG GAG CTG GGG ATT CAG GTA GAT GAA GAC TTT GAA AAC ACT CAG AGA ATT CAA GTT 2206

S S G E N M A G T A E G E G Q Q R H S Q 689  
TCC TCT GGA GAA AAC ATG GCT GGG ACA GCT GAA GGT GAG GGT CAG CAA AGA CAC AGT CAG 2266

L K S S S K S Q A L M P I Q E P G T Q C 709  
CTA AAA AGC TCA TCT AAA AGC CAG GCT CTA ATG CCA ATT CAA GAG CCT GGG ACT CAA TGT 2326

E L S Q N L Q N L Y G T P V F R P V L E 729  
GAG CTC AGC CAG AAT CTT CAG AAT CTC TAT GGT ACC CCA GTA TTC AGG CCT GTT CTA GAG 2386

N S W L F P T R I G G N F N H V S L K A 749  
AAC TCC TGG CTC TTT CCA ACC AGA ATT GGA GGT AAC TTT AAC CAT GTT TCC TTG AAA GCC 2446

S W V M G R P F G S E Q R P K W F H P L 769  
TCC TGG GTT ATG GGC CGC CCC TTT GGG TCA GAG CAG AGG CCT AAG TGG TTC CAT CCT TTG 2506



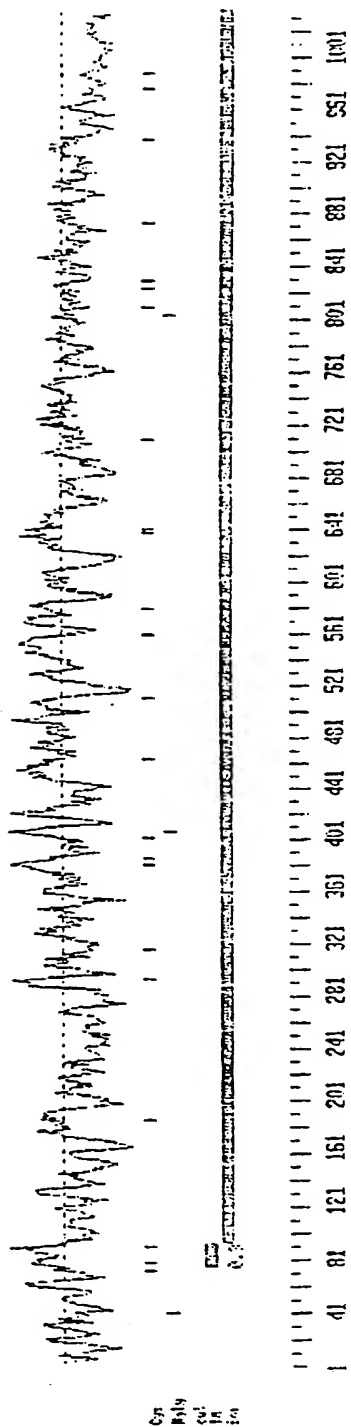
54/58

P F Q N A G A Q G R G K S F G I Q S F H 789  
 CCT TTT CAG AAT GCA GGG GCC CAG GGC CGA GGT AAA AGT TTT GGT ATT CAA TCC TTC CAT 2566  
 P Q I F Y S G E R F M K F S R V A R G C 809  
 CCC CAG ATA TTT TAT TCA GGT GAA AGA TTC ATG AAA TTT TCC AGA GTT GCT CGG GGA TGT 2626  
 H S N G T F G R L P R P I C Q H V Q A C 829  
 CAC TCG AAT GGA ACA TTT GGG AGA CTG CCA AGA CCC ATT TGT CAG CAT GTA CAG GCC TGC 2686  
 P E R P Q M M G T L E R S R A V A S K I 849  
 CCT GAG AGA CCA CAA ATG ATG GGA ACT CTT GAA AGG TCT AGG GCA GTA GCC TCC AAG ATA 2746  
 G H S Y S L D S Q P A R A V G K P W P Q 869  
 GGT CAC TCC TAT TCC CTG GAT TCA CAG CCA GCA AGA GCA GTA GGG AAG CCA TGG CCT CAG 2806  
 Q A C T R V T E L T E A T G K L I R T S 889  
 CAA GCT TGC ACC AGG GTA ACA GAG TTA ACT GAA GCA ACT GGA AAA CTG ATA AGA ACA TCC 2866  
 H I G K P H P Q S F Q P A A A T Q K L R 909  
 CAT ATT GGA AAG CCT CAC CCT CAG TCC TTT CAA CCA GCA GCA GCC ACA CAA AAA CTA AGA 2926  
 P A S Q Q G V Q M K T Q G G A S N P A L 929  
 CCT GCT TCT CAG CAA GGA GTC CAG ATG AAG ACA CAA GGT GGG GCT TCA AAT CCA GCT CTC 2986  
 Q I G S H P M C K S S Q F K S D Q S N P 949  
 CAA ATA GGG TCC CAT CCC ATG TGC AAG AGC TCT CAG TTC AAA TCC GAT CAG TCC AAC CCA 3046  
 S T V K H S Q P K P F H S V P S Q P K S 969  
 TCC ACA GTC AAA CAC TCC CAG CCT AAA CCC TTC CAT TCT GTG CCC TCT CAA CCT AAA TCC 3106  
 S Q T K S C Q S Q P S Q T K P S P C K S 989  
 TCT CAG ACA AAA TCC TGT CAG TCC CAG CCC TCC CAA ACT AAA CCT TCT CCA TGC AAA TCT 3166  
 T Q P K P S Q P W P P Q S K P S Q P R P 1009  
 ACT CAG CCT AAG CCA AGC CAG CCC TGG CCT CCC CAG TCT AAG CCT TCT CAG CCC AGA CCC 3226  
 P Q P K S S S T N P S Q A K A H H S K A 1029  
 CCT CAA CCT AAG TCA TCC TCA ACC AAT CCT TCA CAA GCT AAG GCA CAC CAC TCA AAA GCA 3286  
 G Q K R G G K H \* 1038  
 GGG CAG AAG AGG GGA GGG AAG CAT TAA 3313  
 AGAGCTAACTCCAGAGATCTATAAAGCATATCCTTTACCCAGGCCATTCTCTATCATATAGTAAGCAGAAGAGTTGCCAT 3392  
 GAAAGTAAAGACTACTGTCTATTAGCATGTAAAACAAAGAAAGATATACATGACCGAATTGGATATCTTTGTTTGTGTTG 3471  
 TTTGAGACAGAGTTTCACTCTTGTGTTGCCAGGCTGGAGTGCAATGGCAGATCTCGGCTCACCGCAACCTCTGCTTCCT 3550  
 GGCTTAAAGTGATTCTCTGCCTCAGCCTCTCGAGTAGCTGGGATTACAGGCATGCACCACCACACCCAGCTAATTTTG 3629  
 TATTTTATAGTAGGSCAGGGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACCTCCGACCTCAGGTGATCCGCCCCACCTA 3708  
 GGCTCTCAAAGTGTGGGATTAAGTGTGTAAGCCACAGTGCCCGAGCCGAATTGGATATCTTTAAGATATCTGTAAGT 3787  
 GTTATATCCCTAACCAAGAAGAAAAATATGAAATAATTAAGACTAGAATCAAGCAGTAGATAATTGAATCCAATCTTG 3866  
 GGTATTATTAGATAATGTATAACTTGACCCAGGGAATGGGGTCTATGAGACAACCCCACTTGGAGAAGAATGGGGTT 3945  
 AGGGTCTCTAATTGCAAAGTGACTGTACAATAGGACGAAAGTTGCCTCTGTGTCTGAGAAAGTATCTTAGTTGTTGGCT 4024

55/58

GCTCCAGAGGTATCTTTGTCAAAAGCTTCTGGTTCAATATCAGCCACTGAGCAGATAACCCTGCTTATTTGGTGTGGTT 4103  
AAATCAACTAGCTTCTGCTAATAGCCCCAATTGCTTGAATGGGAAAACCTCTCATTGACCCCTTATAGGTAGAAATA 4182  
ATGAATTAAACAACCAATAAAATTAATCATTGGCATTAAAAAAAAAAAAAAAAAAAAA 4244

FIG. 29  
56/58



CONSENSUS    \*->maederrlLkknrvrliesLgldvLdelLdvLlekdvlnlkeeEkik  
               +++    +    ++ r++l++e+L+ d   d +Ld L +++++ ++e E  
 CARD6        5    STP--SEIERERKKLLEILQHD-PDSILDTLTSRRLISEEEYETLE 48  
 CONSENSUS    ragakledDKarelvdslqrrgsqafdafidaledTgqsyLAdvLel<-\*  
               + l +    r l++ +q++g. + ++ f+ +l++    LA++ +l  
 CARD6        49 NVTDLLKK--SRKLLILVQKKGEATCQHFLKCLFS-TFPQLAAICGL    92

58 / 58

